```
           Copyright   GenCore version 5.1.6
                       Copyright (c) 1993 - 2004 Compugen Ltd.

protein - nucleic search, using frame_plus_p2n model

on:        March 17, 2004, 16:42:09 ; Search time 69 seconds
                                       (without alignments)
                                       651.464 Million cell updates/sec

le:        US-09-648-310-4
fect score: 410
uence:     1 MNVDHEVNLLVEBIHRLGSK.........PGEILLQGVHDDVDIILLQD 81

ring table: BLOSUM62
           Xgapop 10.0 , Xgapext  0.5
           Ygapop 10.0 , Ygapext  0.5
           Fgapop  6.0 , Fgapext  7.0
           Delop   6.0 , Delext   7.0

rched:     682709 seqs, 277475446 residues

al number of hits satisfying chosen parameters:   1365418

imum DB seq length: 0
imum DB seq length: 200000000

t-processing: Minimum Match 0%
              Maximum Match 100%
              Listing first 45 summaries

mand line parameters: -DEV=xlh
DDEL=frame+_p2n.model
/cgn2_1/USPTO_spool/USO9648310/runat_15032004_101255_16593/app_query.fasta_1.263
B=Issued_Patents_NA -QFMT=fastap -SUFFIX=p2n.rni -MINMATCH=0.1 -LOOPCL=0
OOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
ST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
DE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
ER=US09648310 @CGN_1_1_69 @runat_15032004_101255_16593 -NCPU=6 -ICPU=3
) MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
5V_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
GAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

abase :    Issued_Patents_NA:*
           1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
           2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
           3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
           4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
           5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
           6: /cgn2_6/ptodata/2/ina/backfilesl.seq:*
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

sult No.	Score	% Query Match	Length	DB	ID	Description
1	371	90.5	786	4	US-09-023-655-79	Sequence 79, Appl
2	352	85.9	800	4	US-09-644-460-26	Sequence 26, Appl
3	66	16.1	2133	4	US-09-328-352-1774	Sequence 1774, Ap
4	64.5	15.7	393	4	US-09-252-991A-12137	Sequence 12137, A
5	64.5	15.7	2323	1	US-08-358-901-1	Sequence 1, Appli
6	64.5	15.7	2323	1	US-08-566-347-1	Sequence 1, Appli
7	64.5	15.7	2323	1	US-08-693-835-1	Sequence 1, Appli
8	64	15.6	2329	1	US-08-457-797A-8	Sequence 8, Appli
9	64	15.6	2329	1	US-08-812-025-8	Sequence 8, Appli
10	64	15.6	2329	3	US-09-138-873A-8	Sequence 8, Appli
11	63	15.4	1278	4	US-09-134-000C-2571	Sequence 2571, Ap
12	63	15.4	14770	4	US-09-220-132-30	Sequence 30, Appl

13	c	62.5	15.2	675	4	US-09-252-991A-5194	Sequence 5194, Ap
14		62.5	15.2	954	4	US-09-252-991A-5307	Sequence 5307, Ap
15	c	62	15.1	1664976	4	US-08-916-421B-1	Sequence 1, Appli
16		62	14.9	885		US-09-328-352-1423	Sequence 1423, Ap
17		61	14.9	1074	1	US-08-008-216-3	Sequence 3, Appli
18		61	14.9	1074	1	US-08-459-569-3	Sequence 3, Appli
19		61	14.9	1074	1	US-08-458-831-3	Sequence 3, Appli
20		61	14.9	1365	4	US-09-107-532A-1659	Sequence 1659, Ap
21	c	61	14.9	6563	4	US-09-453-702B-192	Sequence 192, App
22		61	14.9	21126	1	US-08-008-216-19	Sequence 19, Appl
23		61	14.9	21126	1	US-08-459-569-19	Sequence 19, Appl
24		61	14.9	21126	1	US-08-458-831-19	Sequence 19, Appl
25		60.5	14.8	2082	4	US-09-107-532A-2158	Sequence 2158, Ap
26	c	60	14.6	582	4	US-09-252-991A-12907	Sequence 12907, A
27	c	60	14.6	1449	4	US-09-252-991A-13298	Sequence 13298, A
28		60	14.6	2028	4	US-09-252-991A-12500	Sequence 12500, A
29		60	14.6	2040	4	US-09-252-991A-12977	Sequence 12977, A
30	c	59.5	14.5	1134	4	US-09-252-991A-8104	Sequence 8104, Ap
31	c	59.5	14.5	1314	4	US-09-252-991A-8001	Sequence 8001, Ap
32	c	59.5	14.5	1755	4	US-09-252-991A-8133	Sequence 8133, Ap
33		59.5	14.5	3003	3	US-08-915-337-1	Sequence 1, Appli
34		59	14.4	1450	4	US-09-554-998A-1	Sequence 1, Appli
35		59	14.4	1740	4	US-08-851-567B-50	Sequence 50, Appl
36		59	14.4	3913	1	US-08-472-934-7	Sequence 7, Appli
37		59	14.4	3913	2	US-08-323-460A-7	Sequence 7, Appli
38		59	14.4	3913	3	US-08-461-146C-7	Sequence 7, Appli
39		59	14.4	3913	3	US-08-461-145C-7	Sequence 7, Appli
40		59	14.4	3913	1	US-08-628-829-11	Sequence 11, Appl
41		59	14.4	4592	9	US-08-472-934-9	Sequence 9, Appli
42		59	14.4	4592	2	US-08-323-460A-9	Sequence 9, Appli
43		59	14.4	4592	2	US-08-461-146C-9	Sequence 9, Appli
44		59	14.4	4592	2	US-08-461-145C-9	Sequence 9, Appli
45		59	14.4	5414	4	US-08-628-829-13	Sequence 13, Appl

ALIGNMENTS

```
RESULT 1
US-09-023-655-79
; Sequence 79, Application US/09023655
; Patent No. 6607879
; GENERAL INFORMATION:
;  APPLICANT:  Cocks, Benjamin G.
;  APPLICANT:  Susan G. Stuart
;  APPLICANT:  Jeffrey J. Seilhamer
;  TITLE OF INVENTION:  COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE
;  TITLE OF INVENTION:  EXPRESSION
;  NUMBER OF SEQUENCES:  1508
;  CORRESPONDENCE ADDRESS:
;   ADDRESSEE:  INCYTE PHARMACEUTICALS, INC.
;   STREET:  3174 PORTER DRIVE
;   CITY:  PALO ALTO
;   STATE:  CALIFORNIA
;   COUNTRY:  USA
;   ZIP:  94304
;  COMPUTER READABLE FORM:
;   MEDIUM TYPE:  Floppy disk
;   COMPUTER:  IBM PC compatible
;   OPERATING SYSTEM:  PC-DOS/MS-DOS
;   SOFTWARE:  Word Perfect 6.1 for Windows/MS-DOS 6.2
;  CURRENT APPLICATION DATA:
;   APPLICATION NUMBER:  US/09/023,655
;   FILING DATE:  HEREWITH
;   CLASSIFICATION:
;  PRIOR APPLICATION DATA:
;   APPLICATION NUMBER:
;   FILING DATE:
;   CLASSIFICATION:
;  ATTORNEY/AGENT INFORMATION:
;   NAME:  Zeller, Karen J.
;   REGISTRATION NUMBER:  37,071
;   REFERENCE/DOCKET NUMBER:  PA-0001 US
;   TELECOMMUNICATION INFORMATION:
```

```
                TELEPHONE:  (650) 855-0555
                TELEFAX:    (650) 845-4166
INFORMATION FOR SEQ ID NO: 79:
     SEQUENCE CHARACTERISTICS:
          LENGTH:  786 base pairs
          TYPE:    nucleic acid
          STRANDEDNESS:  single
          TOPOLOGY: linear
     IMMEDIATE SOURCE:
          LIBRARY:  THP1PLB01
          CLONE:    012364
```

```
Alignment Scores:
Pred. No.:              4.35e-50    Length:        786
Score:                  371.00      Matches:        81
Percent Similarity:     96.43%      Conservative:    0
Best Local Similarity:  96.43%      Mismatches:      0
Query Match:            90.49%      Indels:          3
DB:                                 Gaps:            0

US-09-648-310-4 (1-81) x US-09-023-655-79 (1-786)

Qy   1 MetAsnValAspHisGluValAsnLeuLeuValGluGluIleHisArgLeuGlySerLys 20
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 140 ATGAATGTGGATCACGAGGTTAACCTCTTAGTGGAGGAAATTCATCGTGTTGGGTTCAAA 199

Qy  21 AsnAlaAspGlyLysLeuSer-ValLys-PheGlyValLeuPheArgAspAspLysCysA 40
         |||||||||||||||||    ||||||  |||||||||||||||||||||||||||||
Db 200 AATGCTGATGGAAAGTTAAGCCGTGAAATTTGGGGTCCTCTTCCGTGATGATAAATGTG 259

Qy  40 laAsnLeuPheGluAlaLeuAlaLeuValGlyThrLeuLysAla-LysArgArgLysIleVal 59
         |||||||||||||||||||||||||||||||||||||||||  ||||||||||||||||
Db 260 CCAACCTCTTTGAAGCATTGGTAGGAACTCTTAAAGCTCCAAAAACGAGGAAGATTGTA 319

Qy  60 ThrTyrProGlyGluLeuLeuGlnGlyValHisAspAspValAspIleIleLeuLeuLeu 79
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 320 ACATATCCAGGAGAGAGCTGCTTCTGCAAGGTGTTCATGATGATGTGACATTATTATTACTG 379

Qy  80 GlnAsp 81
         ||||||
Db 380 CAAGAT 385
```

```
ESULT 2
S-09-644-460-26
Sequence 26, Application US/09644460
Patent No. 6657053
GENERAL INFORMATION:
APPLICANT: Fisher, Paul B.
TITLE OF INVENTION: Reciprocal Subtraction Differential
TITLE OF INVENTION: Display
FILE REFERENCE: 34587-C-PCT-USA
CURRENT APPLICATION NUMBER: US/09/644,460
CURRENT FILING DATE: 2000-08-23
PRIOR APPLICATION NUMBER: PCT/US99/04323
PRIOR FILING DATE: 1999-02-26
PRIOR APPLICATION NUMBER: US 09/197,889
PRIOR FILING DATE: 1998-11-23
PRIOR APPLICATION NUMBER: US 09/185,115
PRIOR FILING DATE: 1998-11-03
PRIOR APPLICATION NUMBER: US 09/032,684
PRIOR FILING DATE: 1998-02-27
NUMBER OF SEQ ID NOS: 42
SOFTWARE: FastSEQ for Windows Version 4.0
SEQ ID NO 25
   LENGTH: 800
   TYPE: DNA
   ORGANISM: homo sapiens
US-09-644-460-26

Alignment Scores:
Pred. No.:      5.06e-47   Length:   800
Score:          352.00     Matches:   74
```

```
Percent Similarity:     93.98%      Conservative:    4
Best Local Similarity:  89.16%      Mismatches:      3
Query Match:            85.85%      Indels:          2
DB:                                 Gaps:            0

US-09-648-310-4 (1-81) x US-09-644-460-26 (1-800)

Qy   1 MetAsnValAspHisGluValAsnLeuLeuValGluGluIleHisArgLeuGlySerLys 20
         ||||||||||| ||||||||:||||||||||||||||||||||||||||||||||||||
Db 170 ATGAACGTGGAGCATGAGGTTAACCTCCTGGTGGAGGAAATTCATCGTCTGGGTTCAAA 229

Qy  21 AsnAlaAspGlyLysLeuSerValLysPheGlyValLeuPheArgAspLysCysAla 40
         ||||||||||||||||||||||||||:|||||||||||||||:::|||||||||||
Db 230 AATGGCCGATGGGAAACTGAGTGAGTTTGGGGTCCTCTTCCAAGACGACAGATGTGCC 289

Qy  41 AsnLeuPheGluAla-LeuValGlyThrLeuLysAla-AlaLysArgArgLysIleValT 60
         |||||||||||||  |||||||||||||||||  |||||||||||||||||||||||||
Db 290 AATCTCTTTGAAACCCGTTGGTGGGAACTTGAAAGCCCGCCAAAACGAAGGAAGATTGTTA 349

Qy  60 hrTyrProGlyGluLeuLeuGlnGlyValHisAspAspValAspIleIleLeuLeuLeuG 80
         ||||||||||||||||||||||||||||||||||::|||||||||||||||||||||||
Db 350 CGTACGCAGAGAGAGCTGCTTTTGCAAGGTGTTCATGATGATGTTGACATTGTATTGCTGC 409

Qy  80 lnAsp 81
         |||||
Db 410 AAGAT 414
```

```
RESULT 3
US-09-328-352-1774
; Sequence 1774, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; TITLE OF INVENTION: BAUMANNI FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 1774
; LENGTH: 2133
; TYPE: DNA
; ORGANISM: Acinetobacter baumannii
US-09-328-352-1774

Alignment Scores:
Pred. No.:              1.95        Length:       2133
Score:                  66.00       Matches:        24
Percent Similarity:     50.67%      Conservative:   14
Best Local Similarity:  32.00%      Mismatches:     29
Query Match:            16.10%      Indels:          8
DB:                                 Gaps:            4

US-09-648-310-4 (1-81) x US-09-328-352-1774 (1-2133)

Qy    1 MetAsnValAspHisGluValAsnLeuLeuValGluGluIleHisArgLeuGlySerLys 20
          :||||||||::: |||||||||||||||||||||||||||||||:::|||
Db 1396 ATGGTGGTACGCCGTGAAACAACAAACTGGTGCGTTATGTTCAT---TTAGGAACGGGT 1452

Qy   21 Asn-------AlaAspGlyLysLeuSerValLysPheGlyValLeuPheArgAsp--- 36
          |||       ||:  ::::   :::         ::::::  |||
Db 1453 AACTACCACCATGCTATGAAATGCTCGCATTTATACTGATTACGGCCTAATGACCACCGATAAA 1512

Qy   37 AspLysCysAlaAsn------LeuPheGluAlaLeuLeuValGlyThrLeuLysAlaAla 53
          |||                  :::
Db 1513 GACTTGTGTGAAGACGTACACCGTATTTTCCAAGAGCTCACGGGTAATGGTAAAATGGCA 1572

Qy   54 LysArgArgLysIleValThrTyrProGlyGluLeuLeuGln 68
          |||:::::::::::
Db 1573 AACTGAAAAGTTACTCCATGCACCTTTTACTCTGCATGCTCAG 1617
```

```
RESULT 4
US-09-252-991A-12137
```

Left column

equence 12137, Application US/09252991A
atent No. 6551795
ENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
EQ ID NO 12137
LENGTH: 393
TYPE: DNA
ORGANISM: Pseudomonas aeruginosa
09-252-991A-12137

gnment Scores:
d. No.: 0.293
re: 64.50
cent Similarity: 39.51%
t Local Similarity: 25.93%
ry Match: 15.73%

Length: 393
Matches: 21
Conservative: 11
Mismatches: 38
Indels: 11
Gaps: 2

09-648-310-4 (1-81) x US-09-252-991A-12137 (1-393)

```
  1 MetAsnValAspHisGluValAsnLeuLeuValGluGluIleHisArgLeuGlySerLys 20
    ||||||        :::     :::  ::  ||||||      ||
136 ATGAACGCGACACCAAGGACTCAACCATCCGATCTCCGTGAAGCGGCTG------AAA 189

 21 AsnAlaAspGlyLysLeuSerValLysPheGlyValLeuPheArgAspLysAspLysCysAla 40
    ||||||||        :::  ::       :::::
190 CGCGCCGATGGTCATCTTCGGCACGTCAGTGATGGGATGATCGAAAGTGGCGAATCCTGTCTC 249

 41 AsnLeuPheGluAlaLeuValGlyThrLeuLysAlaAlaLysArgArgIleValThr 60
    :::::::   |||
250 GACATCGCCCGCCAGCTCGCCGGCGTGGAAAAGCGCGGTGACACGGCGCAAAGCGCGTCCTG 309

 61 TyrProGlyGluLeuGluLeuGlnGlyValHisAspAspValAspIleIleLeuLeuGln 80
    :::||||||        :::|:::
310 ------ATCCACGACCACATCGACCATTGCCTGGCCCAT 342

 81 Asp 81
    |||
343 GAC 345
```

ULT 5
-08-358-901-1/c
Sequence 1, Application US/08358901
Patent No. 5554521
GENERAL INFORMATION:
APPLICANT: Suslow, Trevor V.
APPLICANT: Jones, Jonathan D.G.
TITLE OF INVENTION: No. 5554521el Chitinase-Producing Plants
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend Khourie and Crew
STREET: One Market Plaza, Steuart Tower, Suite 2000
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94105
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/358,901

Right column

FILING DATE: 19-DEC-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/550,253
FILING DATE: 09-JUL-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 06/888,033
FILING DATE: 18-JUL-1986
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 06/593,691
FILING DATE: 26-MAR-1984
ATTORNEY/AGENT INFORMATION:
NAME: Smith, William M.
REGISTRATION NUMBER: 30,223
REFERENCE/DOCKET NUMBER: 12176-5-3
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-326-2400
TELEFAX: 415-326-2422
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2323 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 361..2043
US-08-358-901-1

Alignment Scores:
Pred. No.: 3.83
Score: 64.50
Percent Similarity: 52.94%
Best Local Similarity: 29.41%
Query Match: 15.73%
DB: 1

Length: 2323
Matches: 20
Conservative: 16
Mismatches: 31
Indels: 1
Gaps: 1

US-09-648-310-4 (1-81) x US-08-358-901-1 (1-2323)

```
Qy    5 HisGluValAsnLeuLeuValGluGluIleHisArgLeuGlySerLysAsnAlaAspGly 24
        |||     :::     :::  ::: ||||||       :::
Db 1086 CACGCCCTTCTGCGCCTTTTGCAGCGCGGCGAACGGATCGTGGATCGAGATTTTGAA--- 1030

Qy   25 LysLeuSerValLysPheGlyValLeuPheArgAspAspLysCysAlaAsnLeuPheGlu 44
        ||:||        :::             |||      :::
Db 1029 GTCCTCGGCGCCTTGGCAGGAGGCGTTCGAACGCCTGAAGCTGCCTTCAATCTCTTTCAG 970

Qy   45 AlaLeuValGlyThrLeuLysAlaAlaLysArgArgLysIleValThrTyrProGlyGlu 64
        |::|::||                       :::
Db  969 GCTGTCGTTGATGCCATTGCCGCCGAGATCGGGATAAAGCCGTACAGCAGGTGGGTCAG 910

Qy   65 LeuLeuLeuGlnGlyValHisAsp 72
        :::|||        :::  ::
Db  909 GTTTTGCCGCGGATCTTGTCGAC 886
```

RESULT 6
US-08-566-347-1/c
Sequence 1, Application US/08566347
Patent No. 5633450
GENERAL INFORMATION:
APPLICANT: Suslow, Trevor V.
APPLICANT: Jones, Jonathan D.G.
TITLE OF INVENTION: No. 5633450el Chitinase-Producing Plants
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend Khourie and Crew
STREET: One Market Plaza, Steuart Tower, Suite 2000
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94105
COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
  APPLICATION NUMBER: US/08/566,347
  FILING DATE: 01-DEC-1995
  CLASSIFICATION: 800
PRIOR APPLICATION DATA:
  APPLICATION NUMBER: US 08/358,901
  FILING DATE: 19-DEC-1994
PRIOR APPLICATION DATA:
  APPLICATION NUMBER: US 07/550,253
  FILING DATE: 09-JUL-1990
PRIOR APPLICATION DATA:
  APPLICATION NUMBER: US 06/888,033
  FILING DATE: 18-JUL-1986
PRIOR APPLICATION DATA:
  APPLICATION NUMBER: US 06/593,691
  FILING DATE: 26-MAR-1984
ATTORNEY/AGENT INFORMATION:
  NAME: Smith, William M.
  REGISTRATION NUMBER: 30,223
  REFERENCE/DOCKET NUMBER: 12176-5-3
TELECOMMUNICATION INFORMATION:
  TELEPHONE: 415-326-2400
  TELEFAX: 415-326-2422
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
  LENGTH: 2323 base pairs
  TYPE: nucleic acid
  STRANDEDNESS: double
  TOPOLOGY: linear
  MOLECULE TYPE: DNA (genomic)
FEATURE:
  NAME/KEY: CDS
  LOCATION: 361..2043
S-08-566-347-1

Alignment Scores:
red. No.:   3.83
core: 64.50
ercent Similarity: 52.94%
est Local Similarity: 29.41%
uery Match: 15.73%
B: 1

Length: 2323
Matches: 20
Conservative: 16
Mismatches: 31
Indels: 1
Gaps: 1

S-09-648-310-4 (1-81) x US-08-566-347-1 (1-2323)

Qy   5 HisGluValAsnLeuLeuValGluIleHisArgLeuGlySerLysAsnAlaAspGly 24
              || :: ||| ::: ||| :::   :::
Db 1086 CACGCCCTTCTGCGCCTTTGCAGCGGCGGCGAACGGATCGTGGATCGAGATTTTGAA--- 1030

Qy   25 LysLeuSerValLysPheGlyValLeuPheArgAspAspLysCysAlaAsnLeuPheGlu 44
              ||| :::||| || ::: ||||||||
Db 1029 GTCCTCGGCGGCCTTGGCAGGAGCGCTGCAACGCCTGGAAGCTGCCTTCAATCTCTTTCAG 970

Qy   45 AlaLeuValGlyThrLeuLysAlaAlaLysArgLysArgLysIleValThrTyrProGlyGlu 64
              ||||:|||| |||||||
Db   969 GCTGTCGTTGATGCCATTGCCGCCGCAGATCGGGATAAAGCCGTACAGCAGGTGGGTCAG 910

Qy   65 LeuLeuLeuGlnGlyValHisAsp 72
              ||:::|| ::: ||
Db   909 GTTTTGCCGCGGGATCTTGTCGAC 886

ESULT 7
S-08-693-835-1/c
Sequence 1, Application US/08693835
Patent No. 5776448
GENERAL INFORMATION:
APPLICANT: Suslow, Trevor V.
APPLICANT: Jones, Jonathan D.G.
TITLE OF INVENTION: No. 5776448el Chitinase-Producing Plants

NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
  ADDRESSEE: Townsend and Townsend and Crew LLP
  STREET: Two Embarcadero Center, 8th Floor
  CITY: San Francisco
  STATE: California
  COUNTRY: USA
  ZIP: 94111-3834
COMPUTER READABLE FORM:
  MEDIUM TYPE: Floppy disk
  COMPUTER: IBM PC compatible
  OPERATING SYSTEM: PC-DOS/MS-DOS
  SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
  APPLICATION NUMBER: US/08/693,835
  FILING DATE: 01-AUG-1996
  CLASSIFICATION: 435
PRIOR APPLICATION DATA:
  APPLICATION NUMBER: US 08/358,901
  FILING DATE: 19-DEC-1994
PRIOR APPLICATION DATA:
  APPLICATION NUMBER: US 07/550,253
  FILING DATE: 09-JUL-1990
PRIOR APPLICATION DATA:
  APPLICATION NUMBER: US 06/888,033
  FILING DATE: 18-JUL-1986
PRIOR APPLICATION DATA:
  APPLICATION NUMBER: US 06/593,691
  FILING DATE: 26-MAR-1984
ATTORNEY/AGENT INFORMATION:
  NAME: Smith, William M.
  REGISTRATION NUMBER: 30,223
  REFERENCE/DOCKET NUMBER: 12176-5-3
TELECOMMUNICATION INFORMATION:
  TELEPHONE: 415-326-2400
  TELEFAX: 415-326-2422
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
  LENGTH: 2323 base pairs
  TYPE: nucleic acid
  STRANDEDNESS: double
  TOPOLOGY: linear
  MOLECULE TYPE: DNA (genomic)
FEATURE:
  NAME/KEY: CDS
  LOCATION: 361..2043
US-08-693-835-1

Alignment Scores:
Pred. No.:   3.83
Score: 64.50
Percent Similarity: 52.94%
Best Local Similarity: 29.41%
Query Match: 15.73%
DB: 1

Length: 2323
Matches: 20
Conservative: 16
Mismatches: 31
Indels: 1
Gaps: 1

US-09-648-310-4 (1-81) x US-08-693-835-1 (1-2323)

Qy   5 HisGluValAsnLeuLeuValGluIleHisArgLeuGlySerLysAsnAlaAspGly 24
              || :: ||| ::: ||| :::   :::
Db 1086 CACGCCCTTCTGCGCCTTTGCAGCGGCGGCGAACGGATCGTGGATCGAGATTTTGAA--- 1030

Qy   25 LysLeuSerValLysPheGlyValLeuPheArgAspAspLysCysAlaAsnLeuPheGlu 44
              ||| :::||| || ::: ||||||||
Db 1029 GTCCTCGGCGGCCTTGGCAGGAGCGCTGCAACGCCTGGAAGCTGCCTTCAATCTCTTTCAG 970

Qy   45 AlaLeuValGlyThrLeuLysAlaAlaLysArgLeuGlyIleHisArgLeuGlySerLysAsnAlaAspGly 24
              ||||:|||| |||||||
Db 1086 CACGCCCTTCTGCGCCTTTGCAGCGGCGGCGAACGGATCGTGGATCGAGATTTGAA--- 1030

Qy   25 LysLeuSerValLysPheGlyValLeuPheArgAspAspLysCysAlaAsnLeuPheGlu 44
              ||| :::||| || :::
Db 1029 GTCCTCGGCGGCCTTGGCAGGAGCGCTGCAACGCCTGGAAGCTGCCTTCAATCTCTTTCAG 970

Qy   45 AlaLeuValGlyThrLeuLysAlaAlaLysArgLysArgLysIleValThrTyrProGlyGlu 64
              ||||:|||| |||||||
Db   969 GCTGTCGTTGATGCCATTGCCGCCGCAGATCGGGATAAAGCCGTACAGCAGGTGGGTCAG 910

Qy   65 LeuLeuLeuGlnGlyValHisAsp 72
              ||:::|| ::: ||
Db   909 GTTTTGCCGCGGGATCTTGTCGAC 886

SULT 8
-08-457-797A-8/c
Sequence 8, Application US/08457797A
Patent No. 5689045
GENERAL INFORMATION:
APPLICANT:
TITLE OF INVENTION: Transgenic pathogen-resistant organism
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Brumbaugh, Graves, Donohue and Raymond
STREET: 30 Rockefeller Plaza
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10112
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/457,797A
FILING DATE: June 1, 1995
ATTORNEY/AGENT INFORMATION:
REGISTRATION NUMBER: 18,839
REGISTRATION NUMBER: 35,225
REFERENCE/DOCKET NUMBER: A29542 FWC - 37/31335
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 408-2500
TELEFAX: (212) 765-2519
TELEX: 238555
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 2329 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
ORIGINAL SOURCE:
ORGANISM: Serratia marcescens
IMMEDIATE SOURCE:
LIBRARY: Cosmid bank from Serratia marcescens
FEATURE:
NAME/KEY: misc_feature
LOCATION: 1..2329
IDENTIFICATION METHOD: experimental
OTHER INFORMATION: /function= "exo-chitinase"
OTHER INFORMATION: /product= "ChiS protein"
OTHER INFORMATION: /evidence= EXPERIMENTAL
OTHER INFORMATION: /note= "sequence listing of the ChiS gene from a
OTHER INFORMATION: plasmid pLChiS from E.coli A 5187"
3-08-457-797A-8

lignment Scores:
red. No.:   4.63
core: 64.00
ercent Similarity: 55.32%
st Local Similarity: 36.17%
ery Match: 15.61%
3:

Length: 2329
Matches: 17
Conservative: 9
Mismatches: 21
Indels: 0
Gaps: 0

5-09-648-310-4 (1-81) x US-08-457-797A-8 (1-2329)

```
      26 LeuSerValLysPheGlyValLeuPheArgArgAspAspLysCysAlaAlaAsnLeuPheGluAla 45
    1026 CTCGCGGCCCTGGCAGGAGCGCTGCAACGCCTGGAAGCTGCCTTCAATCTCTTTCAGGCT 967

      46 LeuValGlyThrLeuLysAlaAlaLysArgLysIleValThrTyrProGlyGluLeu 65
     966 GTCGTTGATGCCATTGCCCGCCGAGATCGGGATAAAGCCGTACAGCAGGTGGGTCAGGTT 907

      66 LeuLeuGlnGlyValHisAsp 72
```

```
Db    906 TTGGCGCGGGATCTTGTCGAC 886
```

RESULT 9
US-08-812-025-8/c
; Sequence 8, Application US/08812025
; Patent No. 5804184
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: Transgenic pathogen-resistant organism
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Brumbaugh, Graves, Donohue and Raymond
; STREET: 30 Rockefeller Plaza
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10112
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/812,025
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Tenser, Arthur
; REGISTRATION NUMBER: 18,839
; NAME: Kole, Lisa
; REGISTRATION NUMBER: 35,225
; REFERENCE/DOCKET NUMBER: A29542 FWC - 37/31335
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 408-2500
; TELEFAX: (212) 765-2519
; TELEX: 238555
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2329 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; ORIGINAL SOURCE:
; ORGANISM: Serratia marcescens
; IMMEDIATE SOURCE:
; LIBRARY: Cosmid bank from Serratia marcescens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1..2329
; IDENTIFICATION METHOD: experimental
; OTHER INFORMATION: /function= "exo-chitinase"
; OTHER INFORMATION: /product= "ChiS protein"
; OTHER INFORMATION: /evidence= EXPERIMENTAL
; OTHER INFORMATION: /note= "sequence listing of the ChiS gene from a
; OTHER INFORMATION: plasmid pLChiS from E.coli A 5187"
US-08-812-025-8

Alignment Scores:
Pred. No.:   4.63
Score: 64.00
Percent Similarity: 55.32%
Best Local Similarity: 36.17%
Query Match: 15.61%
DB:

Length: 2329
Matches: 17
Conservative: 9
Mismatches: 21
Indels: 0
Gaps: 0

US-09-648-310-4 (1-81) x US-08-812-025-8 (1-2329)

```
Qy    26 LeuSerValLysPheGlyValLeuPheArgArgAspAspLysCysAlaAlaAsnLeuPheGluAla 45
Db  1026 CTCGCGGCCCTGGCAGGAGCGCTGCAACGCCTGGAAGCTGCCTTCAATCTCTTTCAGGCT 967

Qy    46 LeuValGlyThrLeuLysAlaAlaLysArgLysIleValThrTyrProGlyGluLeu 65
```

```
Qy        966 GTCGTTGATGCCATTGCCGCCGCAGATCGGGATAAAGCCGTACAGCAGGTGGGTCAGGTT 907
              :::  |||        :::  |||||    ||||||   |||   |||        |||:::::||
Db        906 TTGCGCCGGATCTTGTCGAC 886
```

RESULT 10
US-09-138-873A-8/c
; Sequence 8, Application US/09138873A
; Patent No. 6271438
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: Transgenic pathogen-resistant organism
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Baker & Botts
; STREET: 30 Rockefeller Plaza
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10112
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/138,873A
; FILING DATE: August 24, 1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Tenser, Arthur
; REGISTRATION NUMBER: 18,839
; NAME: Kole, Lisa
; REGISTRATION NUMBER: 35,225
; REFERENCE/DOCKET NUMBER: A29542 FWC - 37/31335
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 408-2500
; TELEFAX: (212) 765-2519
; TELEX: 238555
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2329 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; ORIGINAL SOURCE:
; ORGANISM: Serratia marcescens
; IMMEDIATE SOURCE:
; LIBRARY: Cosmid bank from Serratia marcescens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1..2329
; IDENTIFICATION METHOD: experimental
; OTHER INFORMATION: /function= "exo-chitinase"
; OTHER INFORMATION: /product= "ChiS protein"
; OTHER INFORMATION: /evidence= EXPERIMENTAL
; OTHER INFORMATION: /note= "sequence listing of the ChiS gene from a
; OTHER INFORMATION: plasmid pLChis from E.coli A 5187"
US-09-138-873A-8

Alignment Scores:
Pred. No.:   4.63   Length: 2329
Score: 64.00   Matches: 17
Percent Similarity: 55.32%   Conservative: 9
Best Local Similarity: 36.17%   Mismatches: 21
Query Match: 15.61%   Indels: 0
DB: Gaps: 0

US-09-648-310-4 (1-81) x US-09-138-873A-8 (1-2329)

Qy   26 LeuSerValLysPheGlyValLeuPheArgAspAspLysCysAlaAsnLeuPheGluAla 45
```

---

```
Db       1026 CTCGCGGCCCTGGCAGGAGCGCTGCAGACGCCTGGAAGCTGCCTTCAATCTCTTTCAGGCT 967
              |||:::  ||        ||| |||   ::: |||||        |||||||::|||  |||:::||
Qy         46 LeuValGlyThrLeuLysAlaAlaLysArgArgLysIleValThrTyrProGlyGluLeu 65
              :::|||   |||  ::: ||||           ::: |||
Db        966 GTCGTTGATGCCATTGCCGCCGCAGATCGGGATAAAGCCGTACAGCAGGTGGGTCAGGTT 907

Qy         66 LeuLeuGlnGlyValHisAsp 72
              |||  ::: ||| ///  |||
Db        906 TTGCGCCGGATCTTGTCGAC 886
```

RESULT 11
US-09-134-000C-2571
; Sequence 2571, Application US/09134000C
; Patent No. 6617156
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; TITLE OF INVENTION: ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 032796-032
; CURRENT APPLICATION NUMBER: US/09/134,000C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/055,778
; PRIOR FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 6812
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2571
; LENGTH: 1278
; TYPE: DNA
; ORGANISM: Enterococcus faecalis
US-09-134-000C-2571

Alignment Scores:
Pred. No.:   2.81   Length: 1278
Score: 63.00   Matches: 21
Percent Similarity: 38.46%   Conservative: 14
Best Local Similarity: 23.08%   Mismatches: 28
Query Match: 15.37%   Indels: 28
DB: Gaps: 2

US-09-648-310-4 (1-81) x US-09-134-000C-2571 (1-1278)

```
Qy          1 MetAsnValAspHisGluValAsnLeuLeuValGluGluIleHisArgLeuLeuGlySerLys 20
              |||   :::   :::    :::   ::: |||||||   ||||||||
Db         55 ATGGACTTGGAAAAGTTTAAACAAAGATCAAAGAGTTGTCAGCAATTGGGAGTGAT 114

Qy         21 AsnAlaAspGlyLysLeuSerVal------------------------------------ 28
              |||  |||                                                :::
Db        115 CCTGCAGGAGGGATGACACCGTTACTTTATACAGATTCGTGGTTGGCAGCACAAAAGTAT 174

Qy         29 ------------------LysPheGlyValLeuPheArgAspAspLysCysAlaAsnLeu 42
                                |||:::||  |||||||||:::|||  |||||||
Db        175 GTGCAAAGCCAAACTGAGGCGTTTGGACTGGAAACGCAATTTGACGAAGTCGGGAATCTG 234

Qy         43 PheGluAlaLeuValGlyThrLeuLysAlaAlaLysArgArgLysIleValThrTyrPro 62
              :::  |||  |||   |||||   :::   ||||||||----------------  :::||
Db        235 TTTTGTCGAGTGGCAGGAACG--------------------GAATTTCCT 264

Qy         63 GlyGluLeuLeuLeuGlnGlyValHisAspAsp 73
              |||||  ||| |||:::|||||
Db        265 CAGGAAACGATTTTAACTGGGTCACATATCGAC 297
```

RESULT 12
US-09-220-132-30
; Sequence 30, Application US/09220132
; Patent No. 6506607
; GENERAL INFORMATION:
; APPLICANT: Shyjan, Andrew W.
; TITLE OF INVENTION: METHODS AND COMPOSITONS FOR THE IDENTIFICATION AND ASSESSMENT
; TITLE OF INVENTION: OF PROSTATE CANCER THERAPIES AND THE DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 07334-074001
; CURRENT APPLICATION NUMBER: US/09/220,132
; CURRENT FILING DATE: 1998-12-23

```
PRIOR APPLICATION NUMBER: US 60/079,303
PRIOR FILING DATE: 1998-03-25
PRIOR APPLICATION NUMBER: US 60/068,821
PRIOR FILING DATE: 1997-12-24
NUMBER OF SEQ ID NOS: 191
SOFTWARE: FastSEQ for Windows Version 4.0
SEQ ID NO 30
LENGTH: 14770
TYPE: DNA
ORGANISM: Homo sapiens
-09-220-132-30

gnment Scores:
ed. No.:                 97.3           Length:         14770
ore:                     63.00          Matches:        27
rcent Similarity:        40.43%         Conservative:   11
st Local Similarity:     28.72%         Mismatches:     26
ery Match:               15.37%         Indels:         30
:                        4              Gaps:           4
```

```
-09-648-310-4 (1-81) x US-09-220-132-30 (1-14770)

    6 GluValAsnLeuLeuValGluGluIleHisArgLeuGlySerLysAsnAlaAspGlyLys 25
      ||| ||  :||   ::::::::||  :::::::::||          ::||      :::||
12523 GAACTGAATTTTCAGTGGATGAAAATCAATCAAATACGTGTGGAAAATCCAAATTCTTTA 12582

   26 LeuSerValLysPheGlyValLeu----------PheArgAspLysCysAlaAsn 41
      :::||         ||| ||:||                ||||:||  || ||
12583 ATTTCTCAGAGCTTCATGTTATTAAAAAAATGGTTACCAGAGACGGAAAAAATGCCACA 12642

   42 LeuPheGluAlaLeuValGlyThrLeuLysAlaAlaLysArgArgLysIleValThrTyr 61
      :::||||||       ||||| ||
12643 ACT--GATGCCTTAACTTCGGTCTTGACAAAATTAATCGAAATAGATATAGTGACA--- 12696

   62 ProGlyGluLeuLeuGlnGly------------------- 69
      ||:|||
12697 --------CTGCTAGAAGGACCAAATATTTGATTATGGAAATATTTCAGGCACCAGA 12744

   70 -------------------ValHisAspValAsp 75
                         ||||||
12745 AGTTTTGCAGATGAGAACAATGTTTTCCATGACCCCTGTTGAT 12786
```

```
SULT 13
-09-252-991A-5194/c
Sequence 5194, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 5194
LENGTH: 675
TYPE: DNA
ORGANISM: Pseudomonas aeruginosa
-09-252-991A-5194

ignment Scores:
ed. No.:                 1.34           Length:         675
ore:                     62.50          Matches:        22
rcent Similarity:        36.05%         Conservative:   9
st Local Similarity:     25.58%         Mismatches:     30
ery Match:               15.24%         Indels:         25
:                        4              Gaps:           3

-09-648-310-4 (1-81) x US-09-252-991A-5194 (1-675)
```

```
    5 HisGluValAsnLeuLeuValGluGluIleHis--------------------- 15
      |||  :::::|||                       ||||| |:::|||
  593 CACAGCATCAGCCTGGGCGGCCGAGGAACTGCACGTTACCCATGGCGGCCGTGAGCCGCAG 534

   16 ----------ArgLeuGlySerLysAsnAlaAspGlyLys 25
                  ||| ||
  533 GTGCGGTTGCTCGAGGACGATCTCGGGGGTGGCCCTGTTCGGCAAGGATGGACGCGGCGTA 474

   26 LeuSerValLysPheGlyValLeuPheArgAspAspLysCysAlaAsnLeuPheGluAla 45
      |||||                             :::    |||||
  473 AAACTCACCGATTCCGGCGTTCGCCTGCGT---GACGCCTGCGGCGATGCGTTCGAGCGA 417

   46 LeuValGlyThrLeuLysAlaAlaLysArgArgLys---------IleValThr 60
      |||                                :::::::
  416 CTGCGTGGCGTCGTGCCGAGCTGGCGGCCGGCAGACCGCCGGAGGCCCGTTCGTCCTCGGC 357

   61 TyrProGlyGluLeuLeu 66
      ||||||
  356 GTACCCGGCAGCCTGCTG 339
```

```
RESULT 14
US-09-252-991A-5307
; Sequence 5307, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 5307
; LENGTH: 954
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-5307

Alignment Scores:
Pred. No.:               2.22           Length:         954
Score:                   62.50          Matches:        22
Percent Similarity:      36.05%         Conservative:   9
Best Local Similarity:   25.58%         Mismatches:     30
Query Match:             15.24%         Indels:         25
DB:                      4              Gaps:           3

US-09-648-310-4 (1-81) x US-09-252-991A-5307 (1-954)

    5 HisGluValAsnLeuLeuValGluGluIleHis--------------------- 15
      |||  :::::|||                       ||||| |:::|||
  127 CACAGCATCAGCCTGGGCGGCCGAGGAACTGCACGTTACCCATGGCGGCCGTGAGCCGCAG 186

   16 ----------ArgLeuGlySerLysAsnAlaAspGlyLys 25
                  ||| ||
  187 GTGCGGTTGCTCGAGGACGATCTCGGGGGTGGCCCTGTTCGGCAAGGATGGACGCGGCGTA 246

   26 LeuSerValLysPheGlyValLeuPheArgAspAspLysCysAlaAsnLeuPheGluAla 45
      |||||                             :::    |||||
  247 AAACTCACCGATTCCGGCGTTCGCCTGCGT---GACGCCTGCGGCGATGCGTTCGAGCGA 303

   46 LeuValGlyThrLeuLysAlaAlaLysArgArgLys---------IleValThr 60
      |||                                :::::::
  304 CTGCGTGGCGTCGTGCCGAGCTGGCGGCCGGCAGACCGCCGGAGGCCCGTTCGTCCTCGGC 363

   61 TyrProGlyGluLeuLeu 66
      ||||||
  364 GTACCCGGCAGCCTGCTG 381
```

```
SULT 15
-08-916-421B-1/c
Sequence 1, Application US/08916421B
Patent No. 6503729
GENERAL INFORMATION:
   TITLE OF INVENTION: Complete Genome Sequence of the Methanogenic Archaeon, Methanococc
   APPLICANT: Bult et al.
   Patent No. 6503729
   TITLE OF INVENTION: jannaschii
   FILE REFERENCE: PB275
   CURRENT APPLICATION NUMBER: US/08/916,421B
   CURRENT FILING DATE: 1997-08-22
   PRIOR APPLICATION NUMBER: US 60/024,428
   PRIOR FILING DATE: 1996-08-22
   NUMBER OF SEQ ID NOS: 3
   SOFTWARE: Patentin version 3.1

SEQ ID NO 1
   LENGTH: 1664976
   TYPE: DNA
   ORGANISM: Methanococcus jannaschii
   FEATURE:
   NAME/KEY: misc_feature
   LOCATION: (28222)..(28222)
   OTHER INFORMATION: n equals a, t, c, or g
   NAME/KEY: misc_feature
   LOCATION: (28257)..(28258)
   OTHER INFORMATION: n equals a, t, c, or g
   NAME/KEY: misc_feature
   LOCATION: (84773)..(84773)
   OTHER INFORMATION: n equals a, t, c, or g
   NAME/KEY: misc_feature
   LOCATION: (84808)..(84808)
   OTHER INFORMATION: n equals a, t, c, or g
   NAME/KEY: misc_feature
   LOCATION: (84812)..(84812)
   OTHER INFORMATION: n equals a, t, c, or g
   NAME/KEY: misc_feature
   LOCATION: (98120)..(98120)
   OTHER INFORMATION: n equals a, t, c, or g
   NAME/KEY: misc_feature
   LOCATION: (98159)..(98159)
   OTHER INFORMATION: n equals a, t, c, or g
   NAME/KEY: misc_feature
   LOCATION: (98239)..(98239)
   OTHER INFORMATION: n equals a, t, c, or g
   NAME/KEY: misc_feature
   LOCATION: (98266)..(98266)
   OTHER INFORMATION: n equals a, t, c, or g
   NAME/KEY: misc_feature
   LOCATION: (98343)..(98343)
   OTHER INFORMATION: n equals a, t, c, or g
   NAME/KEY: misc_feature
   LOCATION: (103998)..(103998)
   OTHER INFORMATION: n equals a, t, c, or g
   NAME/KEY: misc_feature
   LOCATION: (148948)..(148948)
   OTHER INFORMATION: n equals a, t, c, or g
   NAME/KEY: misc_feature
   LOCATION: (163385)..(163385)
   OTHER INFORMATION: n equals a, t, c, or g
   NAME/KEY: misc_feature
   LOCATION: (191989)..(191989)
   OTHER INFORMATION: n equals a, t, c, or g
   NAME/KEY: misc_feature
   LOCATION: (191995)..(191995)
   OTHER INFORMATION: n equals a, t, c, or g
   NAME/KEY: misc_feature
   LOCATION: (231980)..(231980)
   OTHER INFORMATION: n equals a, t, c, or g
   NAME/KEY: misc_feature
   LOCATION: (234187)..(234187)
   OTHER INFORMATION: n equals a, t, c, or g
   NAME/KEY: misc_feature
   LOCATION: (234220)..(234220)
   OTHER INFORMATION: n equals a, t, c, or g
   NAME/KEY: misc_feature
   LOCATION: (234814)..(234814)
   OTHER INFORMATION: n equals a, t, c, or g
   NAME/KEY: misc_feature
   LOCATION: (309398)..(309398)
   OTHER INFORMATION: n equals a, t, c, or g
   NAME/KEY: misc_feature
   LOCATION: (309418)..(309418)
   OTHER INFORMATION: n equals a, t, c, or g
   NAME/KEY: misc_feature
   LOCATION: (312837)..(312837)
   OTHER INFORMATION: n equals a, t, c, or g
   NAME/KEY: misc_feature
   LOCATION: (312993)..(312993)
   OTHER INFORMATION: n equals a, t, c, or g
   NAME/KEY: misc_feature
   LOCATION: (319226)..(319226)
   OTHER INFORMATION: n equals a, t, c, or g
   NAME/KEY: misc_feature
   LOCATION: (559167)..(559167)
   OTHER INFORMATION: n equals a, t, c, or g
   NAME/KEY: misc_feature
   LOCATION: (559241)..(559241)
   OTHER INFORMATION: n equals a, t, c, or g
   NAME/KEY: misc_feature
   LOCATION: (600992)..(600992)
   OTHER INFORMATION: n equals a, t, c, or g
   NAME/KEY: misc_feature
   LOCATION: (622708)..(622708)
   OTHER INFORMATION: n equals a, t, c, or g
   NAME/KEY: misc_feature
   LOCATION: (657081)..(657081)
   OTHER INFORMATION: n equals a, t, c, or g
   NAME/KEY: misc_feature
   LOCATION: (657203)..(657203)
   OTHER INFORMATION: n equals a, t, c, or g
   NAME/KEY: misc_feature
   LOCATION: (674435)..(674435)
   OTHER INFORMATION: n equals a, t, c, or g
   NAME/KEY: misc_feature
   LOCATION: (682442)..(682442)
   OTHER INFORMATION: n equals a, t, c, or g
   NAME/KEY: misc_feature
   LOCATION: (713652)..(713652)
   OTHER INFORMATION: n equals a, t, c, or g
   NAME/KEY: misc_feature
   LOCATION: (741684)..(741684)
   OTHER INFORMATION: n equals a, t, c, or g
   NAME/KEY: misc_feature
   LOCATION: (779455)..(779455)
   OTHER INFORMATION: n equals a, t, c, or g
   NAME/KEY: misc_feature
   LOCATION: (779676)..(779676)
   OTHER INFORMATION: n equals a, t, c, or g
   NAME/KEY: misc_feature
   LOCATION: (855539)..(855539)
   OTHER INFORMATION: n equals a, t, c, or g
   NAME/KEY: misc_feature
   LOCATION: (871619)..(871619)
   OTHER INFORMATION: n equals a, t, c, or g
   NAME/KEY: misc_feature
   LOCATION: (1084830)..(1084830)
   OTHER INFORMATION: n equals a, t, c, or g
   NAME/KEY: misc_feature
   LOCATION: (1096846)..(1096846)
   OTHER INFORMATION: n equals a, t, c, or g
   NAME/KEY: misc_feature
   LOCATION: (1119881)..(1119881)
   OTHER INFORMATION: n equals a, t, c, or g
   NAME/KEY: misc_feature
   LOCATION: (1130881)..(1130881)
```

OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1310988)..(1310988)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1313224)..(1313224)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1349473)..(1349473)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1349491)..(1349491)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1470091)..(1470091)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1569020)..(1569020)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1602912)..(1602912)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1603734)..(1603734)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1637998)..(1637998)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1664854)..(1664855)
OTHER INFORMATION: n equals a, t, c, or g
-08-916-421B-1

ignment Scores:
ed. No.:   1.26e+05   Length: 1664976
ore: 62.00   Matches: 23
rcent Similarity: 47.31%   Conservative: 21
st Local Similarity: 24.73%   Mismatches: 29
ery Match: 15.12%   Indels: 20
;   4   Gaps: 5

-09-648-310-4 (1-81) x US-08-916-421B-1 (1-1664976)

     3 ValAspHisGluValAsn-------LeuLeuValGluGluIleHisArgLeuGlySer 19
       :::::|||:::|:||| ||||| ||| :::||| :::|||
720607 ATTGAGCATAAGGTTAGAGATGCCTTTTGTTGTCAGAAATGCTAAATGCCGTTGGT--- 720551

    20 LysAsnAlaAspGlyLysLeuSerValLysPheGlyValLeuPheArgAspAspLysCys 39
       :::||| :::||| |||::: |||||||:::||
720550 ------AGAAATGGCTTATTGCTGTAGGTATTGGAATCTGCTTAGAAGATGATGAATGC 720497

    40 AlaAsnLeu--------------------PheGluAlaLeuValGlyThrLeuLys 51
       :::   :::||::: |||:::|||
720496 ATAAAGATAGGTAATCAAATCCTATGGGAGTATAAAAAGAACTAATCAATGAACTTAAA 720437

    52 AlaAlaLysArgArgLys--------IleValThrTyrProGlyGluLeuLeuLeuGln 68
       ::: ||| :::|||| |||::: |||||| :::
720436 AGTGTTAAACTAAAAAAATTAAACAACATCATTACTTTGAGGGTAAGAACGGGGATGATT 720377

    69 GlyValHisAspAspValAspIleIleLeuLeuGlnAsp 81
       ||:::   :::||||||::: |||
720376 GGGATT-------AATTGCCCTCTATATTGGTTGATGAT 720347

arch completed: March 17, 2004, 16:46:52
b time : 274 secs

```
       Copyright   GenCore version 5.1.6
       Copyright (c) 1993 - 2004 Compugen Ltd.

protein - nucleic search, using frame_plus_p2n model

 on:   March 17, 2004, 16:42:09 ; Search time 2808 Seconds
                                   (without alignments)
                                   1250.280 Million cell updates/sec

le:       US-09-648-310-4
fect score: 410
quence:   1 MNVDHEVNLLVEEIHRLGSK........PGEILLQGVHDDVDILLQD 81

ring table: BLOSUM62
       Xgapop 10.0 , Xgapext  0.5
       Ygapop 10.0 , Ygapext  0.5
       Fgapop  6.0 , Fgapext  7.0
       Delop   6.0 , Delext   7.0

arched:   3470272 segs, 21671516995 residues

al number of hits satisfying chosen parameters:       6940544

imum DB seq length: 0
imum DB seq length: 2000000000

t-processing: Minimum Match 0%
              Maximum Match 100%
              Listing first 45 summaries

mand line parameters:
ODEL=frame+_p2n.model -DEV=xlh
/cgn2_1/USPTO_spool/US09648310/runat_15032004_101254_16562/app_query.fasta_1.263
3=GenEmbl -QFMT=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
NITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human62 -LIST=45
OCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
UTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
ER=US09648310 @CGN_1_1_3731 @runat_15032004_101254_16562 -NCPU=6 -ICPU=3
MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
EV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
GAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

tabase :   GenEmbl:*
    1:  gb_ba:*
    2:  gb_htg:*
    3:  gb_in:*
    4:  gb_om:*
    5:  gb_ov:*
    6:  gb_pat:*
    7:  gb_ph:*
    8:  gb_pl:*
    9:  gb_pr:*
   10:  gb_ro:*
   11:  gb_sts:*
   12:  gb_sy:*
   13:  gb_un:*
   14:  gb_vi:*
   15:  em_ba:*
   16:  em_fun:*
   17:  em_hum:*
   18:  em_in:*
   19:  em_mu:*
   20:  em_om:*
   21:  em_or:*
   22:  em_ov:*
   23:  em_pat:*
   24:  em_ph:*
   25:  em_pl:*
   26:  em_ro:*
   27:  em_sts:*
   28:  em_un:*
   29:  em_vi:*
   30:  em_htg_hum:*
   31:  em_htg_inv:*
   32:  em_htg_other:*
   33:  em_htg_mus:*
   34:  em_htg_pln:*
   35:  em_htg_rod:*
   36:  em_htg_mam:*
   37:  em_htg_vrt:*
   38:  em_sy:*
   39:  em_htgo_hum:*
   40:  em_htgo_mus:*
   41:  em_htgo_other:*
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	410	100.0	743	9	BC014953	BC014953 Homo sapi
	2	410	100.0	835	6	AX456992	AX456992 Sequence
	3	410	100.0	835	6	AX456995	AX456995 Sequence
	4	410	100.0	876	17	AF116682	Af116682 Homo sapi
	5	390	95.1	526	10	BC039801	BC039801 Mus muscu
	6	390	95.1	780	6	AX456990	AX456990 Sequence
	7	390	95.1	780	6	AX456994	AX456994 Sequence
c	8	378	92.2	170882	10	AC107839	AC107839 Mus muscu
c	9	378	92.2	219661	10	AC117670	AC117670 Mus muscu
c	10	375	91.5	692	10	AF065991	AF065991 Mus muscu
	11	371	90.5	786	6	AR379534	AR379534 Sequence
	12	358	87.3	892	5	AY353856	AY353856 Coturnix
	13	351	85.6	1283	5	AY398399	AY398399 Danio rer
	14	351	85.6	1315	5	BC053269	BC053269 Danio rer
	15	303	73.9	100296	9	AL590308	AL590308 Human DNA
	16	287	70.0	176094	2	AC102256	AC102256 Mus muscu
c	17	287	70.0	198132	2	AC102536	AC102536 Mus muscu
	18	287	70.0	198793	2	AC117858	AC117858 Rattus no
c	19	287	70.0	231507	2	AC137434	AC137434 Rattus no
c	20	287	70.0	239053	2	AC129055	AC129055 Rattus no
	21	286	69.8	486	6	AX505660	AX505660 Sequence
	22	286	69.8	516	8	AY097101	AY097101 Arabidops
	23	276	67.3	543	8	AK062756	AK062756 Oryza sat
	24	273	66.6	239874	5	BX248390	BX248390 Zebrafish
	25	252	61.5	552	6	AX396724	AX396724 Sequence
c	26	227.5	55.5	98124	8	AT16L1	AT16L1 Arabidops
c	27	227.5	55.5	199536	8	ATCHRIV79	ATCHRIV79 Arabidops
c	28	216	52.7	464	8	SL293059	Z93059 S.latifolia
	29	211	51.5	79575	8	AP006357	AP006357 Lotus cor
c	30	210	51.2	150447	2	AC097277	AC097277 Oryza sat
c	31	108	26.3	1285	10	AF336113	AF336113 Rattus no
	32	108	26.3	73341	2	AC136853	AC136853 Rattus no
c	33	108	26.3	220702	2	AC114181	AC114181 Rattus no
	34	107	26.1	174846	10	AC129212	AC129212 Mus muscu
c	35	107	26.1	186137	10	AC099736	AC099736 Mus muscu
	36	104	25.4	1128	10	AF504061	AF504061 Mus muscu
	37	104	25.4	173312	2	BX324201	BX324201 Danio rer
	38	104	25.4	240809	2	BX571763	BX571763 Danio rer
	39	103	25.1	145746	5	AL953858	AL953858 Zebrafish
	40	102	24.9	1146	9	AF503617	AF503617 Homo sapi
	41	102	24.9	1322	6	AX747674	AX747674 Sequence
	42	102	24.9	1322	9	AK092694	AK092694 Homo sapi
	43	102	24.9	2781	9	HSM804735	AL833422 Homo sapi
	44	102	24.9	2832	9	HSM803459	AL832152 Homo sapi
c	45	102	24.9	152130	9	AP000856	AP000856 Homo sapi

ALIGNMENTS

RESULT 1

```
C014953
OCUS        BC014953        743 bp    mRNA    linear   PRI 25-NOV-2003
FINITION    Homo sapiens chromosome 6 open reading frame 115, mRNA (cDNA clone
            MGC:22978 IMAGE:4849571), complete cds.
CCESSION    BC014953
RSION       BC014953.1   GI:15928976
EYWORDS     MGC.
OURCE       Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
FERENCE     1  (bases 1 to 743)
AUTHORS     Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
            Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
            Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
            Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
            Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
            Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
            Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
            Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
            Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
            McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
            Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
            Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
            Fahey,J., Helton,E., Ketteman,M., Madan,A., Rodrigues,S.,
            Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
            Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
            Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
            Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E.,
            Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
TITLE       Generation and initial analysis of more than 15,000 full-length
            human and mouse cDNA sequences
JOURNAL     Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
MEDLINE     22388257
PUBMED      12477992
FERENCE     2  (bases 1 to 743)
AUTHORS     Strausberg,R.
TITLE       Direct Submission
JOURNAL     Submitted (01-OCT-2001) National Institutes of Health, Mammalian
            Gene Collection (MGC), Cancer Genomics Office, National Cancer
            Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
            USA
REMARK      NIH-MGC Project URL: http://mgc.nci.nih.gov
OMMENT      Contact: MGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Dr. Daniel McVicar, DBS/NCI
            cDNA Library Preparation: Rubin Laboratory (LLNL)
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Genome Sequence Centre,
            BC Cancer Agency, Vancouver, BC, Canada
            info@bcgsc.bc.ca
            Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield,
            Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin,
            Letticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo
            Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven
            Ness, Pawan Pandoh, Anna-Liisa Prabhu, Parvaneh Saeedi, Jacqueline
            Schein, Duane Smailus, Michael Smith, Lorraine Spence, Jeff Stott,
            Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy,
            George Yang, Scott Zuyderduyn, Marco Marra.
            Clone distribution: MGC clone distribution information can be found
            through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
            Series: IRAL Plate: 34 Row: a Column: 8.
EATURES     Location/Qualifiers
     source  1..743
             /organism="Homo sapiens"
             /mol_type="mRNA"
             /db_xref="taxon:9606"
             /clone="MGC:22978 IMAGE:4849571"
             /tissue_type="Blood, natural killer cell"
             /clone_lib="NIH_MGC_106"
             /lab_host="DH10B-R"
             /note="Vector: pOTB7"
```

```
     gene    1..743
             /gene="C6orf115"
             /note="synonyms: PRO2013, HSPC280"
             /db_xref="LocusID:58527"

     CDS     74..319
             /codon_start=1
             /product="C6orf115 protein"
             /protein_id="AAH14953.1"
             /db_xref="GI:15928977"
             /db_xref="LocusID:58527"
             /translation="MNVDHEVNLLVEBIHRLGSKNADGKLSVKFGVLFRDDKCANLFE
             ALVGTLKAKRRKIVTYPGELLLQGVHDDVDILLQD"

ORIGIN

Alignment Scores:
Pred. No.:                    1.02e-45        Length:              743
Score:                        410.00          Matches:              81
Percent Similarity:           100.00%         Conservative:          0
Best Local Similarity:        100.00%         Mismatches:            0
Query Match:                  100.00%         Indels:                0
DB:                           9                              Gaps:   0

US-09-648-310-4 (1-81) x BC014953 (1-743)

Qy    1 MetAsnValAspHisGluValAsnLeuLeuValGluGluIleHisArgLeuGlySerLys  20
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   74 ATGAATGTGGATCACGAGGTTAACCTCTTAGTGGAGGAAATTCATCGTTTGGGTTCAAAA 133

Qy   21 AsnAlaAspGlyLysLeuSerValLysPheGlyValLeuPheArgAspAspLysCysAla  40
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  134 AATGCTGATGGAAAGTTAAGCGTGAAATTTGGGGTCCTTCCGTGATGATAAATGTGCC  193

Qy   41 AsnLeuPheGluAlaLeuValGlyThrLeuLysAlaAlaLysArgArgLysIleValThr  60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  194 AACCTCTTTGAAGCATTGGTAGGAACCTTAAAGCTGCAAAACGAAGGAAGATTGTAACA 253

Qy   61 TyrProGlyGluLeuLeuLeuGlnGlyValHisAspAspValAspIleLeuLeuLeuGln  80
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  254 TATCCAGGAGAGCTGCTTCTGCAAGGTGTCATGATGATGTTGACATTATTACTGCAA 313

Qy   81 Asp  81
          |||
Db  314 GAT  316
```

```
RESULT 2
AX456992
LOCUS       AX456992                835 bp    DNA       linear   PAT 06-JUL-2002
DEFINITION  Sequence 3 from Patent WO0216419.
ACCESSION   AX456992
VERSION     AX456992.1   GI:21715784
KEYWORDS    .
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
AUTHORS     Fisher,P.B., Kang,D.C. and Su,Z.Z.
TITLE       Progression suppressed gene 13 (psgen 13) and uses thereof
JOURNAL     Patent: WO 0216419-A 3 28-FEB-2002;
            The Trustees of Columbia University in the City of New York (US)
FEATURES    Location/Qualifiers
     source  1..835
             /organism="Homo sapiens"
             /mol_type="unassigned DNA"
             /db_xref="taxon:9606"

ORIGIN

Alignment Scores:
Pred. No.:                    1.17e-45        Length:              835
Score:                        410.00          Matches:              81
Percent Similarity:           100.00%         Conservative:          0
Best Local Similarity:        100.00%         Mismatches:            0
Query Match:                  100.00%         Indels:                0
```

RESULT 3

```
09-648-310-4 (1-81) x AX456992 (1-835)

  6                          Gaps:        0

  1 MetAsnValAspHisGluValAlaAsnLeuLeuValGluGluIleHisArgLeuGlySerLys  20
    |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
197 ATGAATGTGGATCACGAGGTTACCTCTTAGTGGAGGAAATTCATCGTTTGGGTTCAAAA    256

 21 AsnAlaAspGlyLysLeuSerValLysPheGlyValLeuPheArgAspAspLysCysAla    40
    |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
257 AATGCTGATGGAAAGTTAAGCCTGAAATTTGGGGTCCCTCTTCCGTGATGATAAATGTGCC  316

 41 AsnLeuPheGluAlaLeuValGlyThrLeuLysAlaLysArgLysIleValThr          60
    |||||||||||||||||||||||||||||||||||||||||||||||||||||
317 AACCCTCTTTGAAGCATTGGTAGGAACTCTTAAAGCTGCAAAACGAAGGAAGATTGTAACA  376

 61 TyrProGlyGluLeuLeuLeuGlnGlyValHisAspValAspIleLeuLeuGln          80
    |||||||||||||||||||||||||||||||||||||||||||||||||||||
377 TATCCAGGAGAGCTGCTTCTGCAAGGTGTTCATGATGATGTTGACATTATTACTGCAA    436

 81 Asp  81
    |||
437 GAT  439
```

```
SULT 3
456995
CUS       AX456995     Sequence 6 from Patent WO0216419.
FINITION  AX456995              835 bp    DNA        linear    PAT 06-JUL-2002
CESSION   AX456995     GI:21715786
RSION     AX456995.1   GI:21715786
WORDS
RCE       Homo sapiens (human)
ORGANISM  Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
FERENCE   1
AUTHORS   Fisher,P.B., Kang,D.C. and Su,Z.Z.
TITLE     Progression suppressed gene 13 (psgen 13) and uses thereof
JOURNAL   Patent: WO 0216419-A 6 28-FEB-2002;
          The Trustees of Columbia University in the City of New York (US)
ATURES    Location/Qualifiers
 source   1..835
          /organism="Homo sapiens"
          /mol_type="unassigned DNA"
          /db_xref="taxon:9606"
IGIN

ignment Scores:
ed. No.:                1.17e-45       Length:            835
ore:                    410.00         Matches:            81
rcent Similarity:       100.00%        Conservative:        0
st Local Similarity:    100.00%        Mismatches:          0
ery Match:              100.00%        Indels:              0
:                                      Gaps:                0
```

RESULT 4

```
Qy                 81 Asp  81
                      |||
Db                437 GAT  439

RESULT 4
AF116682
ID   AF116682    standard; mRNA; HUM; 876 BP.
XX
AC   AF116682;
XX
SV   AF116682.1
XX
DT   24-MAY-2000 (Rel. 63, Created)
DT   09-MAY-2001 (Rel. 67, Last updated, Version 2)
XX
DE   Homo sapiens PRO2013 mRNA, complete cds.
XX
KW
XX
OS   Homo sapiens (human)
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia;
OC   Eutheria; Primates; Catarrhini; Hominidae; Homo.
XX
RN   [1]
RP   1-876
RA   Zhang C., Yu Y., Zhang S., Wei H., Zhou G., Ouyang S., Luo L., Bi J.,
RA   Liu M., He F.;
RT   "Functional prediction of the coding sequences of 121 new genes deduced by
RT   analysis of cDNA clones from human fetal liver";
RL   Unpublished.
XX
RN   [2]
RP   1-876
RA   Zhang C., Yu Y., Zhang S., Wei H., Zhou G., Ouyang S., Luo L., Bi J.,
RA   Liu M., He F.;
RT   ;
RL   Submitted (24-DEC-1998) to the EMBL/GenBank/DDBJ databases.
RL   Department of Experimental Hematology, Institute of Radiation Medicine,
RL   Beijing Taiping Road 27, Beijing 100850, P. R. China
XX
DR   SWISS-PROT: Q9P1F3; CFB5_HUMAN.
XX
FH   Key             Location/Qualifiers
FH
FT   source          1..876
FT                   /db_xref="taxon:9606"
FT                   /mol_type="mRNA"
FT                   /organism="Homo sapiens"
FT                   /clone="FLB7547"
FT                   /tissue_type="liver"
FT                   /dev_stage="fetal"
FT   CDS             136..381
FT                   /codon_start=1
FT                   /db_xref="SWISS-PROT:Q9P1F3"
FT                   /note="predicted protein of HQ2013"
FT                   /product="PRO2013"
FT                   /protein_id="AAF71102.1"
FT                   /translation="MNVDHEVNLLVEEIHRLGSKNADGKLSVKFGVLFRDDKCANLFEA
FT                   LVGTLKAKRKIVTYPGELLLQGVHDDVDILLQD"
XX
SQ   Sequence 876 BP; 333 A; 131 C; 167 G; 245 T; 0 other;

Alignment Scores:
Pred. No.:              1.23e-45       Length:            876
Score:                  410.00         Matches:            81
Percent Similarity:     100.00%        Conservative:        0
Best Local Similarity:  100.00%        Mismatches:          0
Query Match:            100.00%        Indels:              0
DB:                                    Gaps:                0

US-09-648-310-4 (1-81) x AF116682 (1-876)

Qy   1 MetAsnValAspHisGluValAlaAsnLeuLeuValGluGluIleHisArgLeuGlySerLys  20
```

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAK Plate: 83 Row: h Column: 12
This clone was selected for full length sequencing because it
passed the following selection criteria: Hexamer frequency ORF
analysis.

FEATURES   Location/Qualifiers
     source   1..526
                   /organism="Mus musculus"
                   /mol_type="mRNA"
                   /strain="C57BL/6J"
                   /db_xref="taxon:10090"
                   /clone="IMAGE:1448067"
                   /tissue_type="Mammary gland"
                   /clone_lib="Soares_mammary_gland_NbMMG"
                   /lab_host="DH10B"
                   /note="Vector: pT7T3D-Pac"
     gene   <1..526
                   /gene="3110003A17Rik"
                   /note="synonym: 1700028I04Rik"
                   /db_xref="LocusID:73112"
                   /db_xref="MGI:1920362"
     CDS   <1..317
                   /gene="3110003A17Rik"
                   /codon_start=3
                   /product="3110003A17Rik protein"
                   /protein_id="AAH39801.1"
                   /db_xref="GI:24980968"
                   /db_xref="LocusID:73112"
                   /db_xref="MGI:1920362"
                   /translation="RVQGSDPRSSSSSVKKEAIGESAMNVEHEVNLLVEEIHRLGSKN
                   ADGKLSVKFGVLFQDDRCANLFEALVGTLKAAKRRKIVTYAGELLLQGVHDDVDIVLL
                   QD"

Alignment Scores:

Pred. No.:   Length: 526
Score: 3.39e-43   Matches: 76
Percent Similarity: 390.00   Conservative: 4
Best Local Similarity: 98.77%   Mismatches: 1
Query Match: 93.83%   Indels: 0
DB: 95.12%   Gaps: 0
                        10

US-09-648-310-4 (1-81) x BC039801 (1-526)

Qy   1 MetAsnValAspHisGluValAsnLeuValGluGluIleHisArgLeuGlySerLys 20
           |||||||||||||||||:::||||||||||||||||||||||||||||||||||
Db   72 ATGAATGTGGAACATGAAGTTAACCTCCTGGTGGAGGAAATTCATCGCCTGGGTTCCAAA 131

Qy   21 AsnAlaAspGlyLysLeuSerValLysPheGlyValLeuPheArgAspAspLysCysAla 40
           |||||||||||||||||||||||||:::||||||||||||||||||||:::||||||
Db   132 AATGCTGATGGAAAATTAAGTGTAAAGTTTGGGGTCCTCTTCCAGGATGACAGATGTGCC 191

Qy   41 AsnLeuPheGluAlaLeuValGlyThrLeuLysAlaAlaLysArgArgLysIleValThr 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   192 AATCTCTTTGAAGCGTTGGTAGGAACTCTGAAAGCTGCAAAACGAAGGAAGATTGTTACA 251

Qy   61 TyrProGlyGluLeuLeuLeuGlnGlyValHisAspAspValIleLeuLeuGlnLeuGln 80
           ||||||||:::|||||||||||||||||:::||||||||||||||||||||||||||
Db   252 TACCCAGGGGAACTACTTTTGCAAGGTGTTCATGATGATGTTGACATTGTATTGCTGCAA 311

Qy   81 Asp 81
           |||
Db   312 GAT 314

RESULT 6
AX456990
LOCUS   AX456990   780 bp   DNA   linear   PAT 06-JUL-2002
DEFINITION   Sequence 1 from Patent WO0216619.
ACCESSION   AX456990
VERSION   AX456990.1 GI:21715783
KEYWORDS

(rotated/second column content)

           |||||||||||||||||||||||||||||||||||||||||||||||||||||
     136 ATGAATGTGGAACATGAAGTTAACCTCCTAGTGGAGGAAATTCATCGTTTGGGTTCAAAA 195

      21 AsnAlaAspGlyLysLeuSerValLysPheGlyValLeuPheArgAspAspLysCysAla 40
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
     196 AATGCTGATGGAAAGTTAAGCGGTGAAATTTGGGGTCCTCTTCCGTGATGATAAATGTGCC 255

      41 AsnLeuPheGluAlaLeuValGlyThrLeuLeuLysAlaAlaLysArgArgLysIleValThr 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||
     256 AACCTCTTTGAAGCATTGGTAGGAACTCTTAAAGCTGCAAAACGAAGGAAGATTGTAACA 315

      61 TyrProGlyGluLeuLeuLeuGlnGlyValHisAspAspValIleLeuLeuGln 80
           ||||||||||||||||||||||||||||||||||||||||||||||||||
     316 TATCCAGGAGAGCTGCTTCTGCAAGGTGTTCATGATGATGTTGACATTATATTACTGCAA 375

      81 Asp 81
           |||
     376 GAT 378

RESULT 5
BC039801
LOCUS   BC039801   526 bp   mRNA   linear   ROD 16-APR-2003
DEFINITION   Mus musculus RIKEN cDNA 3110003A17 gene, mRNA (cDNA clone
             IMAGE:1448067), partial cds.
ACCESSION   BC039801
VERSION   BC039801.1 GI:24980967
KEYWORDS   .
SOURCE   Mus musculus (house mouse)
  ORGANISM   Mus musculus
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1 (bases 1 to 526)
  AUTHORS   Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
             Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
             Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
             Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
             Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
             Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
             Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
             Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
             Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
             McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
             Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
             Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
             Fahey,J., Helton,E., Ketteman,M., Madan,A., Rodrigues,S.,
             Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
             Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
             Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
             Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E.,
             Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
  TITLE   Generation and initial analysis of more than 15,000 full-length
             human and mouse cDNA sequences
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
   MEDLINE   22388257
    PUBMED   12477932
REFERENCE   2 (bases 1 to 526)
  AUTHORS   Strausberg,R.
  TITLE   Direct Submission
  JOURNAL   Submitted (12-NOV-2002) National Institutes of Health, Mammalian
             Gene Collection (MGC), Cancer Genomics Office, National Cancer
             Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
             USA
REMARK   NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT   Contact: MGC help desk
             Email: cgapbbs-r@mail.nih.gov
             Tissue Procurement: Marcello Bento Soares, Ph.D.
             cDNA Library Preparation: M. Bento Soares, University of Iowa
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Institute for Systems Biology
             http://www.systemsbiology.org
             contact: amadan@systemsbiology.org
             Anup Madan, Jessica Fahey, Erin Helton, Mark Ketteman, Anuradha
             Madan, Stephanie Rodrigues, Amy Sanchez and Michelle Whiting

Score: 390.00   Matches: 76
Percent Similarity: 98.77%   Conservative: 4
Best Local Similarity: 93.83%   Mismatches: 1
Query Match: 95.12%   Indels: 0
DB: Gaps: 0

US-09-648-310-4 (1-81) x AX456994 (1-780)

Qy   1 MetAsnValAspHisGluValAsnLeuLeuValGluGluIleHisArgLeuGlySerLys 20
       |||||||||||||||||||||||||||||||||||||||||||:||:::|
Db 170 ATGAACGTGGAGCATGAGGTTAACCTCCTGGTGGAGGAAATTCATCGTCTGGGTTCCAAA 229

Qy 21 AsnAlaAspGlyLysLeuSerValLysPheGlyValLeuPheArgAspLysCysAla 40
Db 230 AATGCCGATGGGAAACTGAGTGTGAAGTTTGGGGTCCTTTTCCAAGACGACAGATGTGCC 289

Qy 41 AsnLeuPheGluAlaLeuValGlyThrLeuLysAlaAlaLysArgArgLysIleValThr 60
Db 290 AATCTCTTTGAAGCGTTGGTGGGAACTCTGAAAGCCGCAAAACCGAAGAAGAATTGTTACG 349

Qy 61 TyrProGlyGluLeuLeuLeuGlnGlyValHisAspAlaSpValAspIleIleLeuLeuGln 80
Db 350 TACGCCAGGAGAGCTGCTTTTGCAAGGTGTTCATGATGATGTTGACATTGTATTGCTGCAA 409

Qy 81 Asp 81
Db 410 GAT 412

RESULT 8
AC107839/c
LOCUS   AC107839   170882 bp   DNA   linear   HTG 24-AUG-2002
DEFINITION Mus musculus clone RP23-284K1, WORKING DRAFT SEQUENCE, 5 ordered
            pieces.
ACCESSION   AC107839
VERSION   AC107839.4 GI:22474886
KEYWORDS   HTG; HTGS_PHASE2; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE   Mus musculus (house mouse)
ORGANISM   Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1 (bases 1 to 170882)
AUTHORS   Birren,B., Nusbaum,C. and Lander,E.
TITLE   Mus musculus, clone RP23-284K1
JOURNAL   Unpublished
REFERENCE   2 (bases 1 to 170882)
AUTHORS   Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
            Anderson,S., Barna,N., Bastien,V., Boguslavkiy,L., Boukhgalter,B.,
            Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B.,
            Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A.,
            Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S.,
            Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,
            Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
            Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
            Kamat,A., Karatas,A., Kells,C., LaRocque,K., Lamazares,R.,
            Landers,T., Lehoczky,J., Levine,R., Liu,G., MacLean,C.,
            Macdonald,P., Major,J., Marquis,N., Matthews,C., McCarthy,M.,
            McEwan,P., McKernan,K., Meldrim,J., Meneus,L., Mihova,T.,
            Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C.,
            Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J.,
            Peterson,K., Phunkhang,P., Pierre,N., Pollara,V., Raymond,C.,
            Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J.,
            Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R., Seaman,S.,
            Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
            Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
            Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
            Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
            Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
TITLE   Direct Submission
JOURNAL   Submitted (24-JAN-2002) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
REFERENCE   3 (bases 1 to 170882)
AUTHORS   Birren,B., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S.,
            Barna,N., Bastien,V., Bloom,T., Boguslavkiy,L., Boukhgalter,B.,

---

URCE   Rattus sp.
ORGANISM   Rattus sp.
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
FERENCE   1
AUTHORS   Fisher,P.B., Kang,D.C. and Su,Z.Z.
TITLE   Progression suppressed gene 13 (psgen 13) and uses thereof
JOURNAL   Patent: WO 0216419-A 1 28-FEB-2002;
            The Trustees of Columbia University in the City of New York (US)
ATURES   Location/Qualifiers
source   1..780
            /organism="Rattus sp."
            /mol_type="unassigned DNA"
            /db_xref="taxon:10118"
IGIN

Alignment Scores:
ed. No.:   5.38e-43   Length: 780
ore: 390.00   Matches: 76
rcent Similarity: 98.77%   Conservative: 4
st Local Similarity: 93.83%   Mismatches: 4
ery Match: 95.12%   Indels: 1
                                    Gaps: 0

-09-648-310-4 (1-81) x AX456990 (1-780)

   1 MetAsnValAspHisGluValAsnLeuLeuValGluGluIleHisArgLeuGlySerLys 20
     |||||||||||||||||||||||||||||||||||||:::||
 170 ATGAACGTGGAGCATGAGGTTAACCTCCTGGTGGAGGAAATTCATCGTCTGGGTTCCAAA 229

  21 AsnAlaAspGlyLysLeuSerValLysPheGlyValLeuPheArgAspLysCysAla 40
 230 AATGCCGATGGGAAACTGAGTGTGAAGTTTGGGGTCCTTTTCCAAGACGACAGATGTGCC 289

  41 AsnLeuPheGluAlaLeuValGlyThrLeuLysAlaAlaLysArgArgLysIleValThr 60
 290 AATCTCTTTGAAGCGTTGGTGGGAACTCTGAAAGCCGCAAAACCGAAGAAGAATTGTTACG 349

  61 TyrProGlyGluLeuLeuLeuGlnGlyValHisAspAlaSpValAspIleIleLeuLeuGln 80
 350 TACGCCAGGAGAGCTGCTTTTGCAAGGTGTTCATGATGATGTTGACATTGTATTGCTGCAA 409

  81 Asp 81
 410 GAT 412

SULT 7
456994
OCUS   AX456994   780 bp   DNA   linear   PAT 06-JUL-2002
DEFINITION Sequence 5 from Patent WO0216419.
CESSION   AX456994
RSION   AX456994.1 GI:21715785
YWORDS   .
OURCE   Rattus sp.
ORGANISM   Rattus sp.
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
EFERENCE   1
AUTHORS   Fisher,P.B., Kang,D.C. and Su,Z.Z.
TITLE   Progression suppressed gene 13 (psgen 13) and uses thereof
JOURNAL   Patent: WO 0216419-A 5 28-FEB-2002;
            The Trustees of Columbia University in the City of New York (US)
EATURES   Location/Qualifiers
source   1..780
            /organism="Rattus sp."
            /mol_type="unassigned DNA"
            /db_xref="taxon:10118"
RIGIN

lignment Scores:
ted. No.:   5.38e-43   Length: 780

Camarata,J., Chang,J., Chazaro,B., Choepel,Y., Collymore,A.,
Cook.A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
Faro,S., Ferreira,P., FitzGerald,M., Gage,D., Galagan,J.,
Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hagos,B.,
Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A.,
Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K.,
Liu,G., MacLean,C., Macdonald,P., Major,J., Matthews,C.,
McCarthy,M., Meldrin,J., Meneus,L., Mihova,T., Mlenga,V.,
Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H.,
O'Connor,T., O'Donnell,P., Pierre,N., Raymond,C., Retta,R., Rise,C., Rogov,P.,
Phunkhang,P., Pierre,N., Raymond,C., Retta,R., Rise,C., Rogov,P.,
Roman,J., Roy,A., Schauer,S., Schupback,R., Seaman,S., Severy,P.,
Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Talamas,J.,
Tesfaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J.,
Zembek,L., Zimmer,A. and Zody,M.

TITLE   Direct Submission
JOURNAL   Submitted (24-AUG-2002) Whitehead Institute/MIT Center for Genome
          Research, 320 Charles Street, Cambridge, MA 02141, USA
DMMENT   On Aug 24, 2002 this sequence version replaced gi:20163124.
          All repeats were identified using RepeatMasker:
          Smit,A.F.A. & Green, P. (1996-1997)
          http://ftp.genome.washington.edu/RM/RepeatMasker.html
          -------------------------- Genome Center
             Center: Whitehead Institute/ MIT Center for Genome Research
             Center code: WIBR
             Web site: http://www-seq.wi.mit.edu
             Contact: sequence_submissions@genome.wi.mit.edu
          -------------------------- Project Information
             Center project name: L20600
             Center clone name: 284_K_1
          -------------------------- Summary Statistics
             Sequencing vector: Plasmid; n/a; 100% of reads
             Chemistry: Dye-terminator Big Dye; 100% of reads
             Assembly program: Phrap; version 0.960731
             Consensus quality: 169045 bases at least Q40
             Consensus quality: 169800 bases at least Q30
             Consensus quality: 169800 bases at least Q20
             Consensus quality: 170139 bases at least Q20
             Insert size: 170000; agarose-fp
             Insert size: 170482; sum-of-contigs
             Quality coverage: 13.6 in Q20 bases; agarose-fp
             Quality coverage: 13.6 in Q20 bases; sum-of-contigs
        * NOTE: This is a 'working draft' sequence. It currently
        * consists of 5 contigs. Gaps between the contigs
        * are represented as runs of N. The order of the pieces
        * is believed to be correct as given, however the sizes
        * of the gaps between them are based on estimates that have
        * provided by the submitter.
        * This sequence will be replaced
        * by the finished sequence as soon as it is available and
        * the accession number will be preserved.
        *   1   107625: contig of 107625 bp in length
        *   107626   107725: gap of 100 bp
        *   107726   111441: contig of 3716 bp in length
        *   111442   111541: gap of 100 bp
        *   111542   120063: contig of 8522 bp in length
        *   120064   120163: gap of 100 bp
        *   120164   161248: contig of 41085 bp in length
        *   161249   161348: gap of 100 bp
        *   161349   170882: contig of 9534 bp in length

FEATURES   Location/Qualifiers
     source   1..170882
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10090"
                     /clone="RP23-284K1"
                     /clone_lib="RPCI-23 Female Mouse BAC"
     misc_feature   1..107625
                     /note="assembly_fragment
                     clone_end:SP6
                     vector_side:left"
     misc_feature   107726..111441

---

     misc_feature   /note="assembly_fragment"
                     11542..120063
     misc_feature   /note="assembly_fragment"
                     120164..161248
     misc_feature   /note="assembly_fragment"
                     161349..170882
                     /note="assembly_fragment
                     clone_end:T7
                     vector_side:right"

ORIGIN

Alignment Scores:
Pred. No.:   1.28e-38   Length: 170882
Score: 378.00   Matches: 74
Percent Similarity: 98.77%   Conservative: 6
Best Local Similarity: 91.36%   Mismatches: 1
Query Match: 92.20%   Indels: 0
DB: Gaps: 0

US-09-648-310-4 (1-81) x AC107839 (1-170882)

Qy   1 MetAsnValAspHisGluValAlaAsnLeuLeuValGluIleHisArgLeuGlySerLys 20
          |||||||||||||||||||||||||||||||||||||||:::|||||||||||||||||||
Db 56551 ATGAATGTGGACATGAAGTTAACCCCTCGTGGAGAAATTAATCGCCTGGGTTCCAAA 56492

Qy   21 AsnAlaAspGlyLyLeuSerValLysPheGlyValLeuPheArgAspAspLysCysAla 40
          ||||||||||||||||||||||||||||:::|||||||||||||||||||||||||||||
Db 56491 AATGCTGATGGAAAATTAAGTGTGAAGTTTGGGGTCCTCTTCCAGAATCACAGATGTGCC 56432

Qy   41 AsnLeuPheGluAlaLeuValGlyThrLeuLysAlaAlaLysArgArgLysIleValThr 60
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 56431 AATCTCTTTGAAGCATTGGTAGGAACTCTGAAAGCTGCAAAACGAAGGAAGATTGTTACA 56372

Qy   61 TyrProGlyGluLeuLeuGlnGlyValHisAspAspValAspIleLeuLeuLeuGln 80
          ||||||||||||||||||||||||||||||||||:::||||||||||||||||||::|||
Db 56371 TACACAGGGGAACTACTTTGCAAGGTGTTCATGATGATGTTGACATTGTATTGCTGCAA 56312

Qy   81 Asp 81
          |||
Db 56311 GAT 56309

RESULT 9
AC117670
LOCUS   AC117670   219661 bp   DNA   linear   ROD 22-OCT-2003
DEFINITION Mus musculus chromosome 7, clone RP23-383P11, complete sequence.
ACCESSION   AC117670
VERSION   AC117670.15 GI:37806507
KEYWORDS   HTG.
SOURCE   Mus musculus (house mouse)
  ORGANISM Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1 (bases 1 to 219661)
  AUTHORS   Birren,B., Nusbaum,C. and Lander,B.
  TITLE   Mus musculus chromosome 7, clone RP23-383P11
  JOURNAL   Unpublished
REFERENCE   2 (bases 1 to 219661)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
            Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavkiy,L.,
            Boukhgalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J.,
            Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Collymore,A.,
            Cook.A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
            Faro,S., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,
            Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
            Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
            Kamat,A., Karatas,A., Kells,C., LaRocque,K., Lamazares,R.,
            Landers,T., Lehcczky,J., Levine,R., Lindblad-Toh,K., Liu,G.,
            MacLean,C., Macdonald,P., Major,J., Marquis,N., Matthews,L.,
            McCarthy,M., McEwan,P., McKernan,P., Meldrim,J., Meneus,L.,
            Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R.,
            Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D.,
            Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,
            Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,

Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R.,
Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

TITLE   Direct Submission
JOURNAL   Submitted (10-APR-2002) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
REFERENCE   3 (bases 1 to 219661)
AUTHORS   Birren,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N.,
            Anderson,M., Arachchi,H.M., Barna,N., Bastien,V., Bloom,T.,
            Boguslavkiy,L., Boukhgalter,B., Camarata,J., Chang,J., Choepel,Y.,
            Collymore,A., Cook,A., Cooke,P., Corum,B., DeArellano,K.,
            Diaz,J.S., Dodge,S., Dooley,K., Dorris,L., Erickson,J., Faro,S.,
            Ferreira,P., FitzGerald,M., Gage,D., Galagan,J., Gardyna,S.,
            Graham,L., Grand-Pierre,N., Hafez,N., Hagopian,D., Hagos,B.,
            Hall,J., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
            Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R.,
            Lindblad-Toh,K., Liu,G., Lui,X., Mabbitt,R., MacLean,C.,
            Macdonald,P., Major,J., Manning,J., Matthews,C., McCarthy,M.,
            Meldrim,J., Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J.,
            Nguyen,C., Nicol,R., Norbu,C., O'Connor,T., O'Donnell,P.,
            Rachupka,A., Ramasamy,U., Raymond,C., Retta,R., Rise,C., Rogov,P.,
            Roman,J., Schauer,S., Schupback,R., Seaman,S., Severy,P., Smith,C.,
            Spencer,B., Stange-Thomann,N., Stubbs,M., Travers,M.,
            Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M.,
            Vassiliev,H., Venkataraman,V.S., Viel,R., Vo,A., Wilson,B., Wu,X.,
            Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
TITLE   Direct Submission
JOURNAL   Submitted (26-SEP-2003) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
REFERENCE   4 (bases 1 to 219661)
AUTHORS   Birren,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N.,
            Anderson,M., Arachchi,H.M., Barna,N., Bastien,V., Bloom,T.,
            Boguslavkiy,L., Boukhgalter,B., Camarata,J., Chang,J., Choepel,Y.,
            Collymore,A., Cook,A., Cooke,P., Corum,B., DeArellano,K.,
            Diaz,J.S., Dodge,S., Dooley,K., Dorris,L., Erickson,J., Faro,S.,
            Ferreira,P., FitzGerald,M., Gage,D., Galagan,J., Gardyna,S.,
            Graham,L., Grand-Pierre,N., Hafez,N., Hagopian,D., Hagos,B.,
            Hall,J., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
            Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R.,
            Lindblad-Toh,K., Liu,G., Lui,X., Mabbitt,R., MacLean,C.,
            Macdonald,P., Major,J., Manning,J., Matthews,C., McCarthy,M.,
            Meldrim,J., Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J.,
            Nguyen,C., Nicol,R., Norbu,C., O'Connor,T., O'Donnell,P.,
            Rachupka,A., Ramasamy,U., Raymond,C., Retta,R., Rise,C., Rogov,P.,
            Roman,J., Schauer,S., Schupback,R., Seaman,S., Severy,P., Smith,C.,
            Spencer,B., Stange-Thomann,N., Stojanovic,N., Stubbs,M.,
            Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M., Wu,X.,
            Vassiliev,H., Venkataraman,V.S., Viel,R., Vo,A., Wilson,B., Wu,X.,
            Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
TITLE   Direct Submission
JOURNAL   Submitted (22-OCT-2003) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT   On Oct 22, 2003 this sequence version replaced gi:35929242.
            All repeats were identified using RepeatMasker:
            Smit, A.F.A. & Green, P. (1996-1997)
            http://ftp.genome.washington.edu/RM/RepeatMasker.html
            ----------- Genome Center
            Center: Whitehead Institute/ MIT Center for Genome Research
            Center code: WIBR
            Web site: http://www-seq.wi.mit.edu
            Contact: sequence_submissions@genome.wi.mit.edu
            ----------- Project Information
            Center project name: L23864
            Center clone name: 383_P_11

FEATURES   Location/Qualifiers
     source   1..219661
                     /organism="Mus musculus"

                                          /mol_type="genomic DNA"
                                          /db_xref="taxon:10090"
                                          /chromosome="7"
                                          /map="7"
                                          /clone="RP23-383P11"
                                          /clone_lib="RPCI-23 Female Mouse BAC"
     misc_feature   1..3981
                     /note="wgs_end_extension
                     clone_end:SP6"
     repeat_region   complement(1..2332)
                     /rpt_family="L1"
     unsure   2108..2116
                     /note="<30 qual SNGL region"
     repeat_region   2348..2731
                     /rpt_family="Lx7"
     repeat_region   2778..3165
                     /rpt_family="MTA"
     repeat_region   3172..3194
                     /rpt_family="AT_rich"
     repeat_region   complement(3197..4215)
                     /rpt_family="Lx2"
     misc_feature   3982..3987
                     /note="clone_boundary
                     clone_end:SP6
                     site:EcoRI"
     repeat_region   4212..5184
                     /rpt_family="Lx2"
     repeat_region   5185..5332
                     /rpt_family="URR1A"
     repeat_region   5370..5913
                     /rpt_family="Lx7"
     repeat_region   5916..6280
                     /rpt_family="ORR1C"
     repeat_region   6304..6373
                     /rpt_family="(GAAA)n"
     repeat_region   6374..6520
                     /rpt_family="B1_MM"
     repeat_region   6521..6560
                     /rpt_family="(GAAA)n"
     repeat_region   6571..6624
                     /rpt_family="(GGAA)n"
     repeat_region   6638..6821
                     /rpt_family="Lx7"
     repeat_region   6816..7029
                     /rpt_family="Lx7"
     repeat_region   complement(7081..7229)
                     /rpt_family="RMER14"
     repeat_region   7492..7693
                     /rpt_family="Lx7"
     repeat_region   7878..7901
                     /rpt_family="(TTAA)n"
     repeat_region   complement(7922..8163)
                     /rpt_family="Lx5"
     repeat_region   complement(8687..8808)
                     /rpt_family="Lx4"
     repeat_region   8822..8862
                     /rpt_family="(TG)n"
     repeat_region   complement(8938..9294)
                     /rpt_family="Lx9"
     repeat_region   complement(9236..9473)
                     /rpt_family="ORR1C"
     repeat_region   9474..9508
                     /rpt_family="(GA)n"
     repeat_region   complement(9509..9654)
                     /rpt_family="ORR1C"
     repeat_region   9680..10488
                     /rpt_family="Lx7"
     repeat_region   complement(11504..11839)
                     /rpt_family="RMER15"
     repeat_region   complement(12240..12595)
                     /rpt_family="ORR1D"
     repeat_region   15784..15874
                     /rpt_family="(TC)n"

```
JOURNAL    Submitted (15-MAY-1998) Anatomy & Cell Biology, SUNY HSC at
           Brooklyn, 450 Clarkson Avenue, Brooklyn, NY 11203, USA
FEATURES             Location/Qualifiers
     source          1..692
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /db_xref="taxon:10090"
                     /tissue_type="thymus"
     CDS             <1..240
                     /codon_start=1
                     /product="unknown"
                     /protein_id="AAL16773.1"
                     /db_xref="GI:16303307"
                     /translation="EFHFFYSINNIFMAAGLQFASCFDRMSSGCRNCLKNGLREQKL
                     FHTSKNVSLPTFPNYLLKKNRFCLFYIVMFCFKV"
ORIGIN

Alignment Scores:
Pred. No.:             4.95e-41        Length:          692
Score:                 375.00          Matches:         74
Percent Similarity:    96.30%          Conservative:    4
Best Local Similarity: 91.36%          Mismatches:      3
Query Match:           91.46%          Indels:          0
DB:                    10              Gaps:            0

US-09-648-310-4 (1-81) x AF065991 (1-692)

Qy    1 MetAsnValAspHisGluValAsnLeuLeuValGluGluIleHisArgLeuGlySerLys  20
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  614 ATGAATGTGGAACATGAAGTTAACCTCCTGGTGGAGGAAATTCATCGCCTGGGTTCCAAA  555

Qy   21 AsnAlaAspGlyLysLeuSerValLysPheGlyValLeuPheArgAspAspLysCysAla  40
        ||||||||||||||||||||||||||||||||||||||||||||| :::||||||
Db  554 AATGCTGATGAAAAATTAAGTGTGAAGTTGGGGGCCTTCTTCCAGGATGACAGATGTGCC  495

Qy   41 AsnLeuPheGluAlaLeuValGlyThrLeuValGlyAlaAlaLysArgArgLysIleValThr  60
        |||||||||||||| |||||||||||||| :::||||||||||||||||||||||
Db  494 AATCTCTTTGAAGCATTGGTAGGAACCTGAAAGCTGCAAAACGACGAAGAATTGTTACA  435

Qy   61 TyrProGlyGluLeuLeuGlnGlyValHisAspAspValAspIleIleLeuLeuLeuGln  80
        |||||||||||||||||||||||||||||||||||||| :::||||||||||
Db  434 TACGCAGGGGAACTACTTTTGCAAGGTGTTCATGATGATGTTGACATTGTATTGCTGCAA  375

Qy   81 Asp  81
        |||
Db  374 GAT  372
```

```
RESULT 11
AR379534
LOCUS       AR379534         786 bp    DNA        linear   PAT 18-DEC-2003
DEFINITION  Sequence 79 from patent US 6607879.
ACCESSION   AR379534
VERSION     AR379534.1  GI:40087168
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 786)
  AUTHORS   Cocks,B.G., Stuart,S.G. and Seilhamer,J.J.
  TITLE     Compositions for the detection of blood cell and immunological
            response gene expression
  JOURNAL   Patent: US 6607879-A 79 19-AUG-2003;
FEATURES             Location/Qualifiers
     source          1..786
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN

Alignment Scores:
Pred. No.:             1.99e-40        Length:          786
Score:                 371.00          Matches:         81
Percent Similarity:    96.43%          Conservative:    0
Best Local Similarity: 96.43%          Mismatches:      0
```

```
repeat_region    complement(19343..19458)
                 /rpt_family="RSINE1"
repeat_region    complement(20062..20271)
                 /rpt_family="B3"
repeat_region    21088..21508
                 /rpt_family="RLTR10C"
repeat_region    21737..21804
                 /rpt_family="(TCTA)n"
repeat_region    complement(22169..22245)
                 /rpt_family="L1M4"
repeat_region    22246..22304
                 /rpt_family="(TCTA)n"
repeat_region    complement(22305..22424)
                 /rpt_family="B1_MM"
repeat_region    complement(22425..22500)
                 /rpt_family="L1M4"
repeat_region    complement(22572..23474)
                 /rpt_family="L1R2_RN"
repeat_region    23474..24237
                 /rpt_family="Lx2B"
repeat_region    complement(24373..25479)
                 /rpt_family="RMER1B"
repeat_region    26828..27029
                 /rpt_family="(TG)n"
```

```
ignment Scores:
ed. No.:                     1.72e-38        Length:          219661
ore:                         378.00          Matches:         74
rcent Similarity:            98.77%          Conservative:    6
st Local Similarity:         91.36%          Mismatches:      1
ery Match:                   92.20%          Indels:          0
                                             Gaps:            0

:-09-648-310-4 (1-81) x AC117670 (1-219661)

,      1 MetAsnValAspHisGluValAsnLeuLeuValGluGluIleHisArgLeuGlySerLys  20
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
  10611  ATGAATGTGGAACATGAAGTTAACCTCCTGGTGGAGGAAATTAATCGCCTGGGTTCCAAA  10670

,     21 AsnAlaAspGlyLysLeuSerValLysPheGlyValLeuPheArgAspAspLysCysAla  40
          |||||||||||||||||||||||||||||||||||||||||| :::||||||||
  10671 AATGCTGATGAAAAATTAAGTGTGAAGTTGGGGGCCTTCTTCCAGGATGACAGATGTGCC  10730

,     41 AsnLeuPheGluAlaLeuValGlyThrLeuValLeuLysAlaAlaLysArgArgLysIleValThr  60
          |||||||||||||||||||||||||| :::||||||||||||||||||||
  10731 AATCTCTTTGAAGCATTGGTAGGAGGAACCTGAAAGCTGCAAAACGACGAAGAATTGTTACA  10790

,     61 TyrProGlyGluLeuLeuGlnGlyValHisAspAspValHisIleLeuLeuLeuLeuGln  80
          ||||||||||||||||||||||||||||| :::||||||||||||||
  10791 TACACAGGGGAACTACTTTGCAAGGTGTTCATGATGATGTTGACATTGTATTGCTGCAA  10850

      81 Asp  81
          ||
  10851 GAT  10853
```

```
ESULT 10
F065991/c
OCUS        AF065991         692 bp    mRNA       linear   ROD 20-OCT-2001
EFINITION   Mus musculus unknown mRNA.
CCESSION    AF065991
ERSION      AF065991.1  GI:16303306
EYWORDS     .
OURCE       Mus musculus (house mouse)
ORGANISM    Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
EFERENCE    1  (bases 1 to 692)
AUTHORS     Ganguly,K., Yang,L.F. and Reddy,P.K.
TITLE       A novel cDNA clone from mouse thymus cDNA library
JOURNAL     Unpublished
EFERENCE    2  (bases 1 to 692)
AUTHORS     Ganguly,K., Yang,L.F. and Reddy,P.K.
TITLE       Direct Submission
```

ry Match:　　　90.49%　　　Indels:　　　3
　　　　　　　　　　6　　　　Gaps:　　　　0

09-648-310-4 (1-81) x AR379534 (1-786)

```
  1 MetAsnValAspHisGluValAsnLeuLeuValGluGluIleHisArgLeuGlySerLys 20
    ATGAATGTGGATCACGAGGTTAACCTCTTAGTGGAGGAAATTCATCGTTTGGGTTCAAAA 199
140 ATGAATGTGGATCACGAGGTTAACCTCTTAGTGGAGGAAATTCATCGTTTGGGTTCAAAA 199
 21 AsnAlaAspGlyLysLeuSer-ValLys-PheGlyValLeuPheArgAspLysCysA 40
    AATGCTGATGGAAAGTTAAGCCGTGAAATTTTGGGGTCCTCTTCCGTGATGATAAATGTG 259
200 AATGCTGATGGAAAGTTAAGCCGTGAAATTTTGGGGTCCTCTTCCGTGATGATAAATGTG 259
 40 lAsnLeuPheGluAlaLeuValGlyThrLeuLysAlaAla-LysArgArgLysIleVal 59
    CCAACCCTTTGAAGCATTGGTAGGAACTCTTAAAGCTGCAAAAACGAAGAAGATTGTA 319
260 CCAACCCTTTGAAGCATTGGTAGGAACTCTTAAAGCTGCAAAAACGAAGAAGATTGTA 319
 60 ThrTyrProGlyLeuLeuLeuGlnGlyValHisAspAspValAspIleIleLeuLeu 79
    ACATATCCAGGAGAGCTGCTTCTGCAAGGTGTTCATGATGATGTTGACATTATTACTG 379
320 ACATATCCAGGAGAGCTGCTTCTGCAAGGTGTTCATGATGATGTTGACATTATTACTG 379
 80 GlnAsp 81
    CAAGAT 385
380 CAAGAT 385
```

ULT 12
53856
US　　AY353856　　　　892 bp　　　mRNA　　　linear　　　VRT 25-AUG-2003
INITION　Coturnix coturnix clone QE8f315 hypothetical protein mRNA, complete
　　　　　cds.
ESSION　AY353856
SION　　AY353856.1　　GI:33946397
WORDS　　.
RCE　　　Coturnix coturnix (common quail)
RGANISM　Coturnix coturnix
　　　　　Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
　　　　　Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
　　　　　Phasianinae; Coturnix.
FERENCE　1 (bases 1 to 892)
AUTHORS　Mott,I.W. and Ivarie,R.D.
TITLE　　cDNA array analysis of lines of Japanese quail divergently selected
　　　　　for four-week body weight
JOURNAL　Unpublished
FERENCE　2 (bases 1 to 892)
AUTHORS　Mott,I.W. and Ivarie,R.D.
TITLE　　Direct Submission
JOURNAL　Submitted (29-JUL-2003) Genetics, University of Georgia, Green
　　　　　Street, Athens, GA 30605, USA
ATURES　　　Location/Qualifiers
　source　　1..892
　　　　　　　　/organism="Coturnix coturnix"
　　　　　　　　/mol_type="mRNA"
　　　　　　　　/db_xref="taxon:9091"
　　　　　　　　/clone="QE8f315"
　　　　　　　　/cell_type="8-day embryo"
　CDS　　　　185..511
　　　　　　　　/codon_start=1
　　　　　　　　/evidence=not experimental
　　　　　　　　/product="hypothetical protein"
　　　　　　　　/protein_id="AAQ55814.1"
　　　　　　　　/db_xref="GI:33946398"
　　　　　　　　/translation="MAGIETCGAGLAPVSSNSREQRWERTTMNVEHEISLLVEEIRRL
　　　　　　　　GTKNADGQVSVKFGVLFADEKCANLFEALVGTLKAAKRKKIVTYQGELLLQGVHDNVD
　　　　　　　　IVLLQD"
IGIN

.ignment Scores:
ed. No.:　　1.31e-38　　　Length:　　892
ore:　　　358.00　　　Matches:　　69
rcent Similarity:　96.19%　　Conservative:　9
st Local Similarity:　85.19%　　Mismatches:　3
ery Match:　　87.32%　　　Indels:　　0
:　　　　5　　　　Gaps:　　　0

US-09-648-310-4 (1-81) x AY353856 (1-892)

```
  1 MetAsnValAspHisGluValAsnLeuLeuValGluGluIleHisArgLeuGlySerLys 20
    ||||||||||||||||||||||||||||||:::::|||||||||||:||||||||||||
266 ATGAACGTGGACCATGAAATAAGCCCTTAGTTGAGGAGATTCGGCGGCTGGGGACCAAA 325
 21 AsnAlaAspGlyLysLeuSerValLysPheGlyValLeuPheArgAspLysCysAla 40
    ||||||||||||||||:|:::|||:||||||||||||||||||||||||||:|||||
326 AATGCCGATGGACAGGTGAGTGTGAAATTGGTGTCCTCGCTGATGAAAGTGTGCC 385
 41 AsnLeuPheGluAlaLeuValGlyThrLeuLysLysAlaAlaLysArgArgLysIleValThr 60
    |||||||||||||||||||||||||||||||||||||::||||||||||||||:
386 AACCTCTTGAGGCCCTGGTGGGAACTCTTAAGGCTGCAAAAGATTGTCACT 445
 61 TyrProGlyLeuLeuLeuGlnGlyValHisAspAspValAspIleIleLeuLeuGln 80
    |||||||||||||||||||||||||:::||||||||||||||:::||||||
446 TATCAAGGAGAGCTGCTTTACAAGGTGTTCATGACAATGTTGATATCGTACTGCAA 505
 81 Asp 81
    |||
506 GAC 508
```

RESULT 13
AY398399
LOCUS　　　AY398399　　　　1283 bp　　mRNA　　linear　　VRT 20-OCT-2003
DEFINITION　Danio rerio clone RK067A2G09 hypothetical protein PRO2013 (PRO2013)
　　　　　mRNA, complete cds.
ACCESSION　AY398399
VERSION　AY398399.1　　GI:37681908
KEYWORDS　.
SOURCE　　Danio rerio (zebrafish)
ORGANISM　Danio rerio
　　　　　Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
　　　　　Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
　　　　　Cypriniformes; Cyprinidae; Danio.
REFERENCE　1 (bases 1 to 1283)
AUTHORS　Song,H.D., Wu,X.Y., Sun,X.J., Zhou,Y., Liu,T.X., Deng,M.,
　　　　　Zhang,G.W., Sheng,Y., Chen,Y., Ruan,Z., Jiang,C.L., Fan,H.Y.,
　　　　　Zon,L.I., Kanki,J.P., Look,A.T. and Chen,Z.
TITLE　　Gene Expression Profiling in the Zebrafish Kidney Marrow Tissue
JOURNAL　Unpublished
REFERENCE　2 (bases 1 to 1283)
AUTHORS　Zhang,G.W., Sun,X.J., Wu,X.Y., Song,H.D., Zhou,Y., Liu,T.X.,
　　　　　Deng,M., Sheng,Y., Chen,Y., Ruan,Z., Jiang,C.L., Fan,H.Y.,
　　　　　Zon,L.I., Kanki,J.P., Look,A.T. and Chen,Z.
TITLE　　Direct Submission
JOURNAL　Submitted (24-SEP-2003) State Key Lab for Medical Genomics,
　　　　　Shanghai Institute of Hematology, Ruijin Hospital Affiliated to
　　　　　Shanghai Second Medical University, 197 Rui Jin Road II, Shanghai
　　　　　200025, P. R. China
FEATURES　　　Location/Qualifiers
　source　　1..1283
　　　　　　　　/organism="Danio rerio"
　　　　　　　　/mol_type="mRNA"
　　　　　　　　/db_xref="taxon:7955"
　　　　　　　　/clone="RK067A2G09"
　　　　　　　　/tissue_type="kidney marrow"
　gene　　　　1..1283
　　　　　　　　/gene="PRO2013"
　CDS　　　　61..306
　　　　　　　　/gene="PRO2013"
　　　　　　　　/codon_start=1
　　　　　　　　/product="hypothetical protein PRO2013"
　　　　　　　　/protein_id="AAQ97832.1"
　　　　　　　　/db_xref="GI:37681909"
　　　　　　　　/translation="MNVEHEVSLLIDEIRRLGSKNADGKTSVKFGVLFNDDQCANLFE
　　　　　　　　ALVGTLKAAKRKKVITFDGELLLQGVHDNVDVLLQD"
ORIGIN

Alignment Scores:
Pred. No.:　　1.77e-37　　　Length:　　1283
Score:　　351.00　　　Matches:　　65

Left column (RESULT 14 region)

```
rcent Similarity:      95.06%      Conservative:   12
st Local Similarity:   80.25%      Mismatches:      4
ery Match:             85.61%      Indels:          0
s:                      5          Gaps:            0
```

-09-648-310-4 (1-81) x AY398399 (1-1283)

```
      1 MetAsnValAspHisGluValAlaAsnLeuLeuValGluGluIleHisArgLeuGlySerLys  20
        ||||||||||||||||||||||||||||:::||||||||||||||||||||||||||||||||
     61 ATGAATGTGGAACATGAAGTCAGTCTGCTCATTGATGAAATCCGCCGACTGGGCAGTAAG    120

     21 AsnAlaAspGlyLysLeuSerValLysPheGlyValLeuPheArgAspAspLysCysAla    40
        ||||||||||||||||||||||||||||||:||:||||||||||||||||||||||||||||
    121 AATGCAGACGGGAAGACCAGCGTCAAATTCGGAGTCTTGTTTAACGATGACCAATGCGCC    180

     41 AsnLeuPheGluAlaAlaLeuValGlyThrLeuLysLysAlaAlaLysArgArgLysIleValThr  60
        |||||||||||||||||||||||||||||||:::||||||||||||||||||||||||||||
    181 AATCTCTTTGAAGCTGCTTTGGTCGGAACACTGAAAGCGGCAAGCGGAGAAGGTGATCACT    240

     61 TyrProGlyGluLeuLeuGlnGlyValHisAspAspValAspIleIleLeuLeuGln      80
        ::::                    ||||||||||||||||||||:::|||
    241 TTCGATGGCGAGCTGCTACTGCAAGGAGTTCACGACCACGTTGATGTCGTATTACTGCAG   300

     81 Asp  81
        |||
    301 GAC  303
```

Middle / References section

BC053269

LOCUS   BC053269   1315 bp   mRNA   linear   VRT 05-JUN-2003
             (bases 1 to 1315)
DEFINITION   Danio rerio cDNA clone IMAGE:6796802, partial cds.
ACCESSION   BC053269
VERSION   BC053269.1   GI:31418949
KEYWORDS
SOURCE   Danio rerio (zebrafish)
  ORGANISM   Danio rerio
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
             Cypriniformes; Cyprinidae; Danio.

REFERENCE   (bases 1 to 1315)
  AUTHORS   Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
             Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
             Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
             Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
             Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
             Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
             Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
             Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
             Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
             McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
             Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
             Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
             Fahey,J., Helton,E., Ketteman,M., Madan,A., Rodrigues,S.,
             Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
             Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
             Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
             Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E.,
             Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
  TITLE   Generation and initial analysis of more than 15,000 full-length
             human and mouse cDNA sequences
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
   MEDLINE   22388257
   PUBMED   12477932
REFERENCE   2 (bases 1 to 1315)
  AUTHORS   Strausberg,R.
  TITLE   Direct Submission
  JOURNAL   Submitted (02-JUN-2003) National Institutes of Health, Mammalian
             Gene Collection (MGC), Cancer Genomics Office, National Cancer
             Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
             USA

REMARK   NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT   Contact: MGC help desk
             Email: cgapbs-r@mail.nih.gov

Right column

Tissue Procurement: Leonard I. Zon, M.D.
cDNA Library Preparation: Invitrogen Corp
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Sequencing Group at the Stanford Human Genome
Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: http://www.shgc.stanford.edu
Contact: (Dickson, Mark) mcd@paxil.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
R. M.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAK Plate: 117 Row: m Column: 14
This clone was selected for full length sequencing because it
passed the following selection criteria: Hexamer frequency ORF
analysis, GenomeScan gene prediction, Similarity but not identity
to protein.

FEATURES   Location/Qualifiers
     source   1..1315
                     /organism="Danio rerio"
                     /mol_type="mRNA"
                     /db_xref="taxon:7955"
                     /clone="IMAGE:6796802"
                     /tissue_type="Kidney, zebrafish"
                     /clone_lib="NCI_CGAP_ZKid1"
                     /lab_host="DH10B"
                     /note="Vector: pCMV-SPORT6.1"

ORIGIN

Alignment Scores:
Pred. No.:   1.83e-37   Length: 1315
Score: 351.00   Matches: 65
Percent Similarity: 95.06%   Conservative: 12
Best Local Similarity: 80.25%   Mismatches: 4
Query Match: 85.61%   Indels: 0
DB: 5   Gaps: 0

US-09-648-310-4 (1-81) x BC053269 (1-1315)

```
Qy    1 MetAsnValAspHisGluValAlaAsnLeuLeuValGluGluIleHisArgLeuGlySerLys  20
        ||||||||||||||||||||||||||||:::||||||||||||||||||||||||||||||||
Db   27 ATGAATGTGGAACATGAAGTCAGTCTGCTCATTGATGAAATCCGCCGACTGGGCAGTAAA    86

Qy   21 AsnAlaAspGlyLysLeuSerValLysPheGlyValLeuPheArgAspAspLysCysAla    40
        ||||||||||||||||||||||||||||||:||:|||||||||||||||:::||||||||||
Db   87 AATGCAGACGGGAAGACCAGCGTCAAATTCGGAGTCTTGTTTAACGATGACCAATGCGCC   146

Qy   41 AsnLeuPheGluAlaAlaLeuValGlyThrLeuLysLysAlaAlaLysArgArgLysIleValThr  60
        ||||||||||||||||||||||||:||:||||||||||||||||||:::|||||||||||||
Db  147 AATCTCTTTGAAGCTGCTTTGGTCGGAACACTGAAGCGGCAAGCGGAAGAAGGTGATCACT   206

Qy   61 TyrProGlyGluLeuLeuGlnGlyValHisAspAspValAspIleIleLeuLeuGln      80
        :::                 ||||||||||||||||||||:::||:::|||||||
Db  207 TTCGATGGCGAGCTGCTACTGCAAGGAGTTCACGACCACAACGTTGATGTCGTATTACTGCAG   266

Qy   81 Asp  81
        |||
Db  267 GAC  269
```

RESULT 15
AL590308

LOCUS   AL590308   100296 bp   DNA   linear   PRI 03-JUL-2001
DEFINITION   Human DNA sequence from clone RP11-501K14 on chromosome 6, complete
             sequence.
ACCESSION   AL590308
VERSION   AL590308.8   GI:11459639
KEYWORDS   HTG.
SOURCE   Homo sapiens (human)
  ORGANISM   Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 100296)
  AUTHORS   Blakey,S.

```
ITLE     Direct Submission
OURNAL   Submitted (03-JUL-2001) Sanger Centre, Hinxton, Cambridgeshire,
         CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk Clone
         requests: clonerequest@sanger.ac.uk
MENT     On Jul 4, 2001 this sequence version replaced gi:14252463.
         During sequence assembly data is compared from overlapping clones.
         Where differences are found these are annotated as variations
         together with a note of the overlapping clone name. Note that the
         variation annotation may not be found in the sequence submission
         corresponding to the overlapping clone, as we submit sequences with
         only a small overlap as described above.
         This sequence was finished as follows unless otherwise noted: all
         regions were either double-stranded or sequenced with an alternate
         chemistry or covered by high quality data (i.e., phred quality >=
         30); an attempt was made to resolve all sequencing problems, such
         as compressions and repeats; all regions were covered by at least
         one plasmid subclone or more than one M13 subclone; and the
         assembly was confirmed by restriction digest. The following
         abbreviations are used to associate primary accession numbers given
         in the feature table with their source databases: Em:, EMBL; Sw:,
         SWISSPROT: Tr:, TREMBL; Wp:, WORMPEP; Information on the WORMPEP
         database can be found at
         http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
         was generated from part of bacterial clone contigs of human
         chromosome 6, constructed by the Sanger Centre Chromosome 6 Mapping
         Group. Further information can be found at
         http://www.sanger.ac.uk/HGP/Chr6
         RP11-501K14 is from the library RPCI-11.2 constructed by the group
         of Pieter de Jong. For further details see
         http://www.chori.org/bacpac/home.htm
         VECTOR: pBACe3.6
         IMPORTANT: This sequence is not the entire insert of clone
         RP11-501K14 It may be shorter because we sequence overlapping
         sections only once, except for a 100 base overlap.
         The true left end of clone RP11-501K14 is at 1 in this sequence.
         The true left end of clone RP1-225E12 is at 98297 in this sequence.
         The true right end of clone RP11-94L3 is at 97304 in this sequence.
ATURES             Location/Qualifiers
     source          1..100296
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="6"
                     /clone="RP11-501K14"
                     /clone_lib="RPCI-11.2"
     repeat_region   2..99
                     /note="MIR repeat: matches 103. .194 of consensus"
     repeat_region   118..199
                     /note="MER47 repeat: matches 2242. .2323 of consensus"
     repeat_region   190..265
                     /note="MER47 repeat: matches 2242. .2317 of consensus"
     repeat_region   268..346
                     /note="MER47 repeat: matches 1. .79 of consensus"
     repeat_region   502..671
                     /note="MIR repeat: matches 7. .185 of consensus"
     repeat_region   1420..1737
                     /note="AluJb repeat: matches 2. .310 of consensus"
     repeat_region   2010..2033
                     /note="12 copies 2 mer tt 100% conserved"
     repeat_region   2650..2961
                     /note="AluSx repeat: matches 3. .306 of consensus"
     repeat_region   5767..5895
                     /note="MIR repeat: matches 21. .166 of consensus"
     repeat_region   6062..6159
                     /note="L2 repeat: matches 2587. .2709 of consensus"
     repeat_region   6165..6268
                     /note="L2 repeat: matches 2607. .2709 of consensus"
     repeat_region   6916..7157
                     /note="AluJb repeat: matches 1. .295 of consensus"
     repeat_region   7301..7442
                     /note="L2 repeat: matches 1768. .1916 of consensus"
     repeat_region   7746..7999
                     /note="AluSx repeat: matches 45. .298 of consensus"
     repeat_region   8013..8152
                     /note="MIR repeat: matches 14. .151 of consensus"
     repeat_region   8899..8959
                     /note="MIR repeat: matches 194. .256 of consensus"
     repeat_region   9209..9497
                     /note="L1M88 repeat: matches 5863. .6175 of consensus"
     repeat_region   9503..9698
                     /note="Aluo repeat: matches 90. .281 of consensus"
     repeat_region   9703..10428
                     /note="L1M88 repeat: matches 5126. .5867 of consensus"
     repeat_region   10429..10759
                     /note="AluJb repeat: matches 1. .310 of consensus"
     repeat_region   10760..10773
                     /note="L1M88 repeat: matches 5114. .5126 of consensus"
     repeat_region   10897..11050
                     /note="L2 repeat: matches 821. .962 of consensus"
     repeat_region   11051..11360
                     /note="AluSg repeat: matches 1. .310 of consensus"
     repeat_region   11361..12691
                     /note="L2 repeat: matches 962. .2419 of consensus"
     repeat_region   12850..13009
                     /note="AluJb repeat: matches 129. .279 of consensus"
     repeat_region   15295..15341
                     /note="MIR repeat: matches 97. .142 of consensus"
     repeat_region   15342..15632
                     /note="AluSg repeat: matches 21. .308 of consensus"
     repeat_region   15633..15735
                     /note="MIR repeat: matches 142. .250 of consensus"
     repeat_region   16229..16532
                     /note="AluSx repeat: matches 1. .301 of consensus"
     repeat_region   16577..16796
                     /note="MER20 repeat: matches 1. .217 of consensus"
     repeat_region   18247..18407
                     /note="MER58C repeat: matches 40. .88 of consensus"
     repeat_region   18586..18687
                     /note="MIR repeat: matches 48. .149 of consensus"
     misc_feature    19512..20572
                     /note="CpG island"
                     /evidence=not_experimental
     misc_feature    20013..20256
                     /note="Sequence from overlapping clone RP11-94L3
                     (AL591033). Assembly confirmed by restriction digest."
     repeat_region   20025..20120
                     /note="4 copies 24 mer 72% conserved"
     repeat_region   20058..20117
                     /note="20 copies 3 mer gcg 73% conserved"
     repeat_region   21404..21522
                     /note="MER46C repeat: matches 82. .191 of consensus"
     repeat_region   21843..22026
                     /note="L2 repeat: matches 2351. .2547 of consensus"
     repeat_region   22027..22320
                     /note="AluJb repeat: matches 1. .297 of consensus"
     repeat_region   22321..22369
                     /note="L2 repeat: matches 2547. .2593 of consensus"
     repeat_region   22394..22607
                     /note="MIR repeat: matches 32. .247 of consensus"
     repeat_region   22674..22981
                     /note="AluJb repeat: matches 3. .311 of consensus"
     repeat_region   22982..23029
                     /note="6 copies 8 mer gaaagaaa 79% conserved"
     repeat_region   23030..23106
                     /note="L2 repeat: matches 2667. .2749 of consensus"
     repeat_region   23146..23213
                     /note="MER6 repeat: matches 799. .865 of consensus"
     repeat_region   23216..23317
                     /note="LTR38 repeat: matches 3. .98 of consensus"
     repeat_region   23306..23808
                     /note="LTR38 repeat: matches 46. .611 of consensus"
     repeat_region   23810..24586
                     /note="MER6 repeat: matches 2. .802 of consensus"
     repeat_region   24777..25093
                     /note="AluSx repeat: matches 1. .312 of consensus"
     repeat_region   25107..25470
```

```
repeat_region      /note="THE1C repeat: matches 1. .370 of consensus"
                   25474. .27051
                   /note="THE1C-internal repeat: matches 5. .1580 of
                   consensus"
repeat_region      27052. .27421
                   /note="THE1C repeat: matches 1. .369 of consensus"
repeat_region      27806. .28116
                   /note="AluY repeat: matches 1. .309 of consensus"
repeat_region      28722. .29399
                   /note="L1MD2 repeat: matches 5302. .6111 of consensus"
repeat_region      29400. .29708
                   /note="AluSx repeat: matches 1. .310 of consensus"
repeat_region      29709. .29861
                   /note="L1MD2 repeat: matches 5164. .5302 of consensus"
repeat_region      30862. .31350
                   /note="HERVL repeat: matches 4611. .5143 of consensus"
repeat_region      31360. .31887
                   /note="MER31B repeat: matches 2. .477 of consensus"
repeat_region      32357. .32689
                   /note="LTR16A repeat: matches 91. .444 of consensus"
repeat_region      32948. .33222
                   /note="AluSx repeat: matches 1. .286 of consensus"
repeat_region      33717. .34028
                   /note="AluY repeat: matches 1. .310 of consensus"
repeat_region      34911. .35200
                   /note="AluSp repeat: matches 6. .302 of consensus"
repeat_region      35384. .35496
                   /note="MIR repeat: matches 32. .147 of consensus"
repeat_region      36388. .36686
                   /note="AluSx repeat: matches 1. .302 of consensus"
repeat_region      37165. .37220
                   /note="7 copies 8 mer tccttcct 75% conserved"
repeat_region      37227. .37266
                   /note="MIR repeat: matches 211. .250 of consensus"
repeat_region      38122. .38421
                   /note="AluSx repeat: matches 11. .310 of consensus"
repeat_region      38644. .38737
                   /note="L2 repeat: matches 2616. .2717 of consensus"
repeat_region      39325. .39632
                   /note="Aluo repeat: matches 4. .310 of consensus"
repeat_region      40283. .40636
                   /note="THE1A repeat: matches 1. .353 of consensus"
```

```
lignment Scores:
ed. No.:                   9.04e-29          Length:          100296
core:                      303.00            Matches:         60
rcent Similarity:          100.00%           Conservative:    1
st Local Similarity:       98.36%            Mismatches:      0
ery Match:                 73.90%            Indels:          0
3:                         9                 Gaps:            0
```

```
-09-648-310-4 (1-81) x AL590308 (1-100296)

       21 AsnAlaAspGlyLysLeuSerValLysPheGlyValLeuPheArgAspAspLysCysAla  40
          ::: |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
   74868 GATGCTGATGGAAAGTTAAGCGTGAAATTGGGGTCCTTCCTTCCGTGATGAAAATGTGCC  74927

       41 AsnLeuPheGluAlaLeuValGlyThrLeuLysAlaAlaLysArgArgLysIleValThr  60
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
   74928 AACCTCTTTGAAGCATTGGTAGGAACTCTTAAAGCTGCAAAACGAAGGAAGATTGTAACA  74987

       61 TyrProGlyGluLeuLeuGlnGlyValHisAspAspValAspIleIleLeuLeuGln  80
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
   74988 TATCCAGGAGGAGCTGCTTCTGCAAGGTGTTCATGATGATGTTGACATTATTACTGCAA  75047

       81 Asp  81
          |||
   75048 GAT  75050
```

```
           GenCore version 5.1.6
      Copyright (c) 1993 - 2004 Compugen Ltd.

protein - nucleic search, using frame_plus_p2n model

  on:   March 17, 2004, 16:42:09 ; Search time 1241 Seconds
                                    (without alignments)
                                    240.298 Million cell updates/sec

le:         US-09-648-310-4
fect score: 410
quence:     1 MNVDHEVNLLVEEIHRLGSK.........PGEILLQGVHDDVDIILLQD 81

oring table: BLOSUM62
      Xgapop  10.0 , Xgapext   0.5
      Ygapop  10.0 , Ygapext   0.5
      Fgapop   6.0 , Fgapext   7.0
      Delop    6.0 , Delext    7.0

st-processing: Minimum Match 0%
               Maximum Match 100%
               Listing first 45 summaries

rched:    242557 seqs, 1840798884 residues

al number of hits satisfying chosen parameters:    4865114

imum DB seq length: 0
timum DB seq length: 2000000000

mand line parameters: -DEV=xlh
DEL=frame+_p2n.model -DEV=xlh
=/cgn2_1/USPTO_spool/US09648310/runat_15032004_101255_16630/app_query.fasta_1.263
B=Published Applications NA -QFMT=fastap -SUFFIX=p2n.rnpb -MINMATCH=0.1
OPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62
ANS=human40.cdi -LIST=45 -MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0
R_MIN=0 -ALIGN=15 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100
AXLEN=2000000000 -USER=US09648310 @CGN 1 1 390 @runat 15032004 101255 16630
PU=6 -ICPU=3 -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100
NGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
GAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

abase :   Published Applications NA:*
     1:  /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
     2:  /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
     3:  /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
     4:  /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
     5:  /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
     6:  /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
     7:  /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
     8:  /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
     9:  /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
    10:  /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
    11:  /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
    12:  /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
    13:  /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
    14:  /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
    15:  /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
    16:  /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
    17:  /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
    18:  /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

```
  %
sult  Query
No.  Score  Match Length DB  ID                          Description
-----------------------------------------------------------------------------------
```

No.	% Score	Match	Length	Pred No.	ID	Description
1	100.0	410	410	12	US-10-085-783A-21733	Sequence 21733, A
2	100.0	410	400	15	US-10-242-535A-21733	Sequence 21733, A
3	100.0	410	483	15	US-10-085-783A-56189	Sequence 56189, A
4	100.0	410	483	15	US-10-242-535A-56189	Sequence 56189, A
5	94.4	387	778	9	US-09-925-300-545	Sequence 545, App
6	72.0	295	458	12	US-09-918-095-26075	Sequence 26075, A
7	69.8	286	533	12	US-10-425-114-19867	Sequence 19867, A
8	69.8	286	486	9	US-09-938-842A-355	Sequence 355, App
9	69.8	286	486	11	US-09-938-842A-355	Sequence 355, App
10	65.6	269	492	10	US-09-770-961-618	Sequence 618, App
11	61.5	252	716	12	US-10-424-599-94878	Sequence 94878, A
12	61.5	252	552	13	US-09-920-300A-939	Sequence 939, App
13	61.5	252	552	14	US-10-033-528-939	Sequence 939, App
14	56.3	231	552	9	US-10-099-926-939	Sequence 939, App
15	52.7	216	268	9	US-09-294-093B-2020	Sequence 2020, Ap
16	52.7	216	388	10	US-09-925-299-318	Sequence 318, App
17	36.3	149	388	9	US-09-294-093B-4150	Sequence 4150, Ap
18	26.3	108	388	15	US-10-191-803-820	Sequence 820, App
19	25.6	105	60	15	US-09-908-975-22102	Sequence 22102, A
20	24.9	102	1322	15	US-10-104-047-1199	Sequence 1199, Ap
21	21.0	86	175	9	US-09-294-093B-5438	Sequence 5438, Ap
22	17.4	71.5	1107	13	US-10-425-114-33343	Sequence 33343, A
23	17.4	71.5	2163	12	US-10-425-114-30828	Sequence 30828, A
24	17.1	70	1728	12	US-10-282-122A-16022	Sequence 16022, A
25	16.3	67	2062	12	US-10-424-599-142059	Sequence 142059,
26	16.1	66	1257	12	US-10-282-122A-14898	Sequence 14898, A
27	16.1	66	201143	12	US-10-240-425-1099	Sequence 1099, Ap
28	15.7	65	983	12	US-10-425-114-29078	Sequence 29078, A
29	15.7	64.5	1395	15	US-10-369-493-41656	Sequence 41656, A
30	15.7	64.5	3812	15	US-10-104-047-981	Sequence 981, App
31	15.6	64	527	14	US-10-027-632-283900	Sequence 283900,
32	15.6	64	1338	12	US-10-282-122A-31488	Sequence 31488, A
33	15.6	64	1338	12	US-10-282-122A-33407	Sequence 33407, A
34	15.4	63.5	1395	15	US-10-369-493-44316	Sequence 44316, A
35	15.4	63	1692	14	US-10-148-724A-2	Sequence 2, Appli
36	15.4	63	1692	14	US-10-148-724A-5	Sequence 5, Appli
37	15.4	63	1692	14	US-10-148-724A-10	Sequence 10, Appl
38	15.4	63	2358	14	US-10-104-047-1025	Sequence 1025, Ap
39	15.4	63	5638	15	US-10-369-493-27209	Sequence 27209, A
40	15.4	63	6207	14	US-10-205-194-118	Sequence 118, App
41	15.4	63	14770	10	US-09-873-367C-230	Sequence 230, App
42	15.4	63	32768	9	US-09-070-927A-123	Sequence 123, APP
43	15.2	62.5	526	15	US-10-027-632-289080	Sequence 289080,
44	15.2	62.5	597	15	US-10-027-632-267226	Sequence 267226,

ALIGNMENTS

```
RESULT 1
US-10-085-783A-21733
; Sequence 21733, Application US/10085783A
; Publication No. US2004003784IA1
; GENERAL INFORMATION:
;  APPLICANT: ChondroGene Inc.
;  APPLICANT: Liew, C.C.
;  TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis
;  FILE REFERENCE: 4231/2002
;  CURRENT APPLICATION NUMBER: US/10/085,783A
;  CURRENT FILING DATE:   2002-02-28
;  PRIOR APPLICATION NUMBER: US 60/305,340
;  PRIOR FILING DATE: 2001-07-13
;  PRIOR APPLICATION NUMBER: US 60/275,017
;  PRIOR FILING DATE: 2001-03-12
;  PRIOR APPLICATION NUMBER: US 60/271,955
;  PRIOR FILING DATE: 2001-02-28
;  NUMBER OF SEQ ID NOS: 58994
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 21733
;  LENGTH: 400
;  TYPE: DNA
;  ORGANISM: Human
;  FEATURE:
```

NAME/KEY: misc_feature
LOCATION: (23)..(23)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc_feature
LOCATION: (48)..(48)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc_feature
LOCATION: (328)..(328)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
US-10-242-535A-21733

Alignment Scores:
Pred. No.:   2.82e-58   Length: 400
Score: 410.00   Matches: 81
Percent Similarity: 100.00%   Conservative: 0
Best Local Similarity: 100.00%   Mismatches: 0
Query Match: 100.00%   Indels: 0
DB: 15   Gaps: 0

US-09-648-310-4 (1-81) x US-10-242-535A-21733 (1-400)

Qy   1 MetAsnValAspHisGluValAsnLeuLeuValGluGluIleHisArgLeuGlySerLys 20
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   77 ATGAATGTGGATCACGAGGTTAACCTCTTAGTGGAGGAAATTCATCGTTTGGGTTCAAAA 136

Qy   21 AsnAlaAspGlyLysLeuSerValLysPheGlyValLeuPheArgAspAspLysCysAla 40
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   137 AATGCTGATGGAAAGTTAAGCGTGAAATTTGGGGTCCTCTTCCGTGATGATAAAATGTGCC 196

Qy   41 AsnLeuPheGluAlaLeuValGlyThrLeuLysAlaAlaLysArgArgLysIleValThr 60
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   197 AACCTCTTTGAAGCATTGGTAGGAACTCTTAAAGCTGCAAAACGAAGGAAGATTGTAACA 256

Qy   61 TyrProGlyLeuLeuLeuGlnGlyValHisAspAspValAspIleIleLeuLeuLeuGln 80
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   257 TATCCAGGAGAGCTGCTTCTGCAAGGTGTTCATGATGATGTTGACATTATTATTACTGCAA 316

Qy   81 Asp 81
           |||
Db   317 GAT 319

RESULT 3
US-10-085-783A-56189
; Sequence 56189, Application US/10085783A
; Publication No. US20040037841A1
; GENERAL INFORMATION:
; APPLICANT: ChondroGene Inc.
; APPLICANT: Liew, C.C.
; TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis
; FILE REFERENCE: 4231/2002
; CURRENT APPLICATION NUMBER: US/10/085,783A
; CURRENT FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: US 60/305,340
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/275,017
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: US 60/271,955
; PRIOR FILING DATE: 2001-02-28
; NUMBER OF SEQ ID NOS: 58994
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 56189
; LENGTH: 483
; TYPE: DNA
; ORGANISM: Human
US-10-085-783A-56189

Alignment Scores:
Pred. No.:   3.73e-58   Length: 483
Score: 410.00   Matches: 81
Percent Similarity: 100.00%   Conservative: 0
Best Local Similarity: 100.00%   Mismatches: 0
Query Match: 100.00%   Indels: 0
DB: 12   Gaps: 0

US-09-648-310-4 (1-81) x US-10-085-783A-56189 (1-483)

; NAME/KEY: misc_feature
; LOCATION: (23)..(23)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (48)..(48)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (328)..(328)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; US-10-242-535A-21733

Alignment Scores:
Pred. No.:   2.82e-58   Length: 400
Score: 410.00   Matches: 81
Percent Similarity: 100.00%   Conservative: 0
Best Local Similarity: 100.00%   Mismatches: 0
Query Match: 100.00%   Indels: 0
DB: 15   Gaps: 0

US-09-648-310-4 (1-81) x US-10-242-535A-21733 (1-400)

Qy   1 MetAsnValAspHisGluValAsnLeuLeuValGluGluIleHisArgLeuGlySerLys 20
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   77 ATGAATGTGGATCACGAGGTTAACCTCTTAGTGGAGGAAATTCATCGTTTGGGTTCAAAA 136

Qy   21 AsnAlaAspGlyLysLeuSerValLysPheGlyValLeuPheArgAspAspLysCysAla 40
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   137 AATGCTGATGGAAAGTTAAGCGTGAAATTTGGGGTCCTCTTCCGTGATGATAAAATGTGCC 196

Qy   41 AsnLeuPheGluAlaLeuValGlyThrLeuLysAlaAlaLysArgArgLysIleValThr 60
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   197 AACCTCTTTGAAGCATTGGTAGGAACTCTTAAAGCTGCAAAACGAAGGAAGATTGTAACA 256

Qy   61 TyrProGlyLeuLeuLeuGlnGlyValHisAspAspValAspIleIleLeuLeuLeuGln 80
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   257 TATCCAGGAGAGCTGCTTCTGCAAGGTGTTCATGATGATGTTGACATTATTATTACTGCAA 316

Qy   81 Asp 81
           |||
Db   317 GAT 319

RESULT 2
US-10-242-535A-21733
; Sequence 21733, Application US/10242535A
; Publication No. US20040013663A1
; GENERAL INFORMATION:
; APPLICANT: ChondroGene Inc.
; APPLICANT: Liew, C.C.
; TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis
; FILE REFERENCE: 4231/2005
; CURRENT APPLICATION NUMBER: US/10/242,535A
; CURRENT FILING DATE: 2002-09-12
; PRIOR APPLICATION NUMBER: US 10/085,783
; PRIOR FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: US 60/305,340
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/275,017
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: US 60/271,955
; PRIOR FILING DATE: 2001-02-28
; NUMBER OF SEQ ID NOS: 58994
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 21733
; LENGTH: 400
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (23)..(23)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:

Qy   81 Asp 81

Db   321 GAT 323

RESULT 5
US-09-925-300-545
; Sequence 545, Application US/09925300
; Patent No. US2002015168lA1
; GENERAL INFORMATION:
; APPLICANT: Craig Rosen,
; APPLICANT: Steve Ruben
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA101
; CURRENT APPLICATION NUMBER: US/09/925,300
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05988
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1890
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 545
; LENGTH: 778
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (641)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (652)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-925-300-545

Alignment Scores:
Pred. No.:   7.58e-58   Length: 778
Score: 410.00   Matches: 81
Percent Similarity: 100.00%   Conservative: 0
Best Local Similarity: 100.00%   Mismatches: 0
Query Match: 100.00%   Indels: 0
DB: Gaps: 0

US-09-648-310-4 (1-81) x US-09-925-300-545 (1-778)

Qy   1 MetAsnValAspHisGluValAsnLeuLeuValGluGluIleHisArgLeuGlySerLys 20
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   186 ATGAATGTGGATCACGAGGTTAACCTCTTAGTGGAGGAAATTCATCGTTTGGGTTCAAAA 245

Qy   21 AsnAlaAspGlyLysLeuSerValLysPheGlyValLeuPheArgAspAspLysCysAla 40
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   246 AATGCTGATGGAAAGTTAAGCGTGAAATTTGGGGTCCTTCCGTGATGATAAAATGTGCC 305

Qy   41 AsnLeuPheGluAlaLeuValGlyThrLeuLysAlaAlaLysArgArgLysIleValThr 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   306 AACCTCTTTGAAGCATTGGTAGGAACTCTTAAAGCTGCAAAACGAAGGAAGATTGTAACA 365

Qy   61 TyrProGlyGluLeuLeuGlnGlyValHisAspAspValAspIleIleLeuLeuGln 80
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   366 TATCCAGGAGAGCTGCTTCTGCAAGGTGTTCATGATGATGTTGACATTATATTACTGCAA 425

Qy   81 Asp 81
           |||
Db   426 GAT 428

RESULT 6
US-09-918-995-26075
; Sequence 26075, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; TITLE OF INVENTION: FROM VARIOUS cDNA LIBRARIES

Qy   1 MetAsnValAspHisGluValAsnLeuLeuValGluGluIleHisArgLeuGlySerLys 20
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   81 ATGAATGTGGATCACGAGGTTAACCTCTTAGTGGAGGAAATTCATCGTTTGGGTTCAAAA 140

Qy   21 AsnAlaAspGlyLysLeuSerValLysPheGlyValLeuPheArgAspAspLysCysAla 40
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   141 AATGCTGATGGAAAGTTAAGCGTGAAATTTGGGGTCCTTCCGTGATGATAAAATGTGCC 200

Qy   41 AsnLeuPheGluAlaLeuValGlyThrLeuLysAlaAlaLysArgArgLysIleValThr 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   201 AACCTCTTTGAAGCATTGGTAGGAACTCTTAAAGCTGCAAAACGAAGGAAGATTGTAACA 260

Qy   61 TyrProGlyGluLeuLeuGlnGlyValHisAspAspValAspIleIleLeuLeuGln 80
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   261 TATCCAGGAGAGCTGCTTCTGCAAGGTGTTCATGATGATGTTGACATTATATTACTGCAA 320

Qy   81 Asp 81
           |||
Db   321 GAT 323

RESULT 4
US-10-242-535A-56189
; Sequence 56189, Application US/10242535A
; Publication No. US20040013663A1
; GENERAL INFORMATION:
; APPLICANT: ChondroGene Inc.
; APPLICANT: Liew, C.C.
; TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis
; FILE REFERENCE: 4231/2005
; CURRENT APPLICATION NUMBER: US/10/242,535A
; CURRENT FILING DATE: 2002-09-12
; PRIOR APPLICATION NUMBER: US 10/085,783
; PRIOR FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: US 60/305,340
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/275,017
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: US 60/271,955
; PRIOR FILING DATE: 2001-02-28
; NUMBER OF SEQ ID NOS: 58994
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 56189
; LENGTH: 483
; TYPE: DNA
; ORGANISM: Human
US-10-242-535A-56189

Alignment Scores:
Pred. No.:   3.73e-58   Length: 483
Score: 410.00   Matches: 81
Percent Similarity: 100.00%   Conservative: 0
Best Local Similarity: 100.00%   Mismatches: 0
Query Match: 100.00%   Indels: 0
DB: Gaps: 15

US-09-648-310-4 (1-81) x US-10-242-535A-56189 (1-483)

```
FILE REFERENCE: 20411-756
CURRENT APPLICATION NUMBER: US/09/918,995
CURRENT FILING DATE: 2001-07-30
PRIOR APPLICATION NUMBER: US/09/235,076
PRIOR FILING DATE: 1999-01-20
NUMBER OF SEQ ID NOS: 38054
SOFTWARE: FastSEQ for Windows Version 3.0
SEQ ID NO 26075
  LENGTH: 458
  TYPE: DNA
  ORGANISM: Homo sapiens
  FEATURE:
  NAME/KEY: misc_feature
  LOCATION: (1)...(458)
  OTHER INFORMATION: n = A,T,C or G
US-09-918-995-26075

Alignment Scores:
Pred. No.:           2.27e-54      Length:         458
Score:               387.00        Matches:        76
Percent Similarity:  100.00%       Conservative:   0
Best Local Similarity: 100.00%     Mismatches:     0
Query Match:         94.39%        Indels:         0
DB:                                Gaps:           0

US-09-648-310-4 (1-81) x US-09-918-995-26075 (1-458)

Qy    1 MetAsnValAspHisGluValAsnLeuValGluGluIleHisArgLeuGlySerLys 20
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  231 ATGAATGTGGATCACGAGGTTAACCTCTTAGTGGAGGAAATTCATCGTTTGGGTTCAAAA 290

Qy   21 AsnAlaAspGlyLysLeuSerValLysPheGlyValLeuPheArgAspAspLysCysAla 40
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  291 AATGCTGATGGAAAAGTTAAGCGTGAAAATTTGGGGTCCGTCCTTCCGTGATGATAAAGTGCC 350

Qy   41 AsnLeuPheGluAlaLeuVaLGlyThrLeuLysAlaAlaLysArgArgLysIleValThr 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  351 AACCTCTTTGAGGCATTGGTAGGAACTCTTAAAGCTGCAAAACGAAGGAAGATTGTAACA 410

Qy   61 TyrProGlyGluLeuLeuLeuGlnGlyValHisAspAspValAspIle 76
           |||||||||||||||||||||||||||||||||||||||||||||||
Db  411 TATCCAGGAGAGCTGCTTCTGCAAGGTGTTCATGATGTTGACATT 458
```

```
RESULT 7
US-10-425-114-19867
  Sequence 19867, Application US/10425114
  Publication No. US20040034888A1
GENERAL INFORMATION:
  APPLICANT: Liu, Jingdong
  APPLICANT: Zhou, Yihua
  APPLICANT: Kovalic, David K.
  APPLICANT: Screen, Steven E
  APPLICANT: Tabaska, Jack E
  APPLICANT: Cao, Yongwei
  TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
  TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
  FILE REFERENCE: 38-21(53313)B
  CURRENT APPLICATION NUMBER: US/10/425,114
  CURRENT FILING DATE: 2003-04-28
  NUMBER OF SEQ ID NOS: 73128
  SEQ ID NO 19867
  LENGTH: 533
  TYPE: DNA
  ORGANISM: Zea mays
  FEATURE:
  OTHER INFORMATION: Clone ID: LIB3153-013-B3_FLI
US-10-425-114-19867

Alignment Scores:
Pred. No.:           5.37e-39      Length:         533
Score:               295.00        Matches:        57
Percent Similarity:  83.54%        Conservative:   9
Best Local Similarity: 72.15%      Mismatches:     13
```

```
Query Match:         71.95%        Indels:         0
DB:                                Gaps:           0

US-09-648-310-4 (1-81) x US-10-425-114-19867 (1-533)

Qy    1 MetAsnValAspHisGluValAsnLeuValGluGluIleHisArgLeuGlySerLys 20
           |||||||||||||| |:::
Db   66 ATGAACGTGGAGAGGAGGTCGGGAGGCTCAAGGAGGAGATCCATAGGCTCGGCCAGCAG 125

Qy   21 AsnAlaAspGlyLysLeuSerValLysPheGlyValLeuPheArgAspAspLysCysAla 40
           ||||||
Db  126 CAGCCCGATGGCTCTTACAAGGTCAAGTTGGCGTCCTTCTTCAACGACGACCGGTGTGCA 185

Qy   41 AsnLeuPheGluAlaLeuAlaLeuValGlyThrLeuLysAlaAlaLysArgArgLysIleValThr 60
           ||:::|||| |||:::|||| |::||
Db  186 AATATCTTTGAGCACTAGTTGCACCTTGAGGGCCGCAAGAGAGGAAGATCTTGACC 245

Qy   61 TyrProGlyGluLeuLeuLeuGlnGlyValHisAspAspValAspIleIleLeuLeu 79
           | |:::||||||:::::|||| |||:::|:::|
Db  246 TACGAAGGCGAGCTGCTTCTGCAAGGTGTTCCATGACAACGTGGAGGATAACCCGTGTTG 302
```

```
RESULT 8
US-09-938-842A-355
  Sequence 355, Application US/09938842A
  Patent No. US20020160378A1
GENERAL INFORMATION:
  APPLICANT: Harper, Jeff
  APPLICANT: Kreps, Joel
  APPLICANT: Wang, Xun
  APPLICANT: Zhu, Tong
  TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
  TITLE OF INVENTION: SAME, AND METHODS OF USE
  FILE REFERENCE: SCRIP1300-3
  CURRENT APPLICATION NUMBER: US/09/938,842A
  CURRENT FILING DATE: 2001-08-24
  PRIOR APPLICATION NUMBER: US 60/227,866
  PRIOR FILING DATE: 2000-08-24
  PRIOR APPLICATION NUMBER: US 60/264,647
  PRIOR FILING DATE: 2001-01-16
  PRIOR APPLICATION NUMBER: US 60/300,111
  PRIOR FILING DATE: 2001-06-22
  NUMBER OF SEQ ID NOS: 5379
  SEQ ID NO 355
  LENGTH: 486
  TYPE: DNA
  ORGANISM: Arabidopsis thaliana
US-09-938-842A-355

Alignment Scores:
Pred. No.:           1.46e-37      Length:         486
Score:               286.00        Matches:        55
Percent Similarity:  80.77%        Conservative:   8
Best Local Similarity: 70.51%      Mismatches:     15
Query Match:         69.76%        Indels:         0
DB:                                Gaps:           0

US-09-648-310-4 (1-81) x US-09-938-842A-355 (1-486)

Qy    1 MetAsnValAspHisGluValAsnLeuValGluGluIleHisArgLeuGlySerLys 20
           |||||||||:::|||
Db  199 ATGAACGTAGACGAAGAGATTCAGAACTGGAAGAAGATCCATCGTCTTGGTTCTCGT 258

Qy   21 AsnAlaAspGlyLysLeuSerValLysPheGlyValLeuPheArgAspAspLysCysAla 40
           |||:::||||| |||:::||
Db  259 CAGAACGACGACGGCTCTTACAAGGTGACGTTTGGAGTGTTGTTCAATGACGATGATGTGCC 318

Qy   41 AsnLeuPheGluAlaLeuAlaLeuValGlyThrLeuLysAlaAlaLysArgArgLysIleValThr 60
           |||:::|||| |
Db  319 AACATCTTTGAAGCATTGGTTGGGACTCTGAGAGCTGCCAAGAAACGCCAAGAAAATAGTCGCA 378

Qy   61 TyrProGlyGluLeuLeuLeuGlnGlyLysValHisAspAspValAspIleIleLeu 78
           |:::|||||| |||:::|||
Db  379 TTTGAGGTGAACTTTTGCTTCAGGCGTTCACGATAAGGTTGAGATCACTCTC 432
```

ULT 9
09-938-842A-355
equence 355, Application US/09938842A
ublication No. US20040009476A9
ENERAL INFORMATION:
, APPLICANT: Harper, Jeff
, APPLICANT: Kreps, Joel
, APPLICANT: Wang, Xun
, APPLICANT: Zhu, Tong
, TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
, TITLE OF INVENTION: SAME, AND METHODS OF USE
, FILE REFERENCE: SCRIP1300-3
, CURRENT APPLICATION NUMBER: US/09/938,842A
, CURRENT FILING DATE: 2001-08-24
, PRIOR APPLICATION NUMBER: US 60/227,866
, PRIOR FILING DATE: 2000-08-24
, PRIOR APPLICATION NUMBER: US 60/264,647
, PRIOR FILING DATE: 2001-01-16
, PRIOR APPLICATION NUMBER: US 60/300,111
, PRIOR FILING DATE: 2001-06-22
, NUMBER OF SEQ ID NOS: 5379
, SEQ ID NO 355
,   LENGTH: 486
,   TYPE: DNA
,   ORGANISM: Arabidopsis thaliana
09-938-842A-355

gnment Scores:
d. No.:   1.46e-37   Length: 486
re: 286.00   Matches: 55
cent Similarity: 80.77%   Conservative: 8
t Local Similarity: 70.51%   Mismatches: 15
ry Match: 69.76%   Indels: 0
                        11   Gaps: 0

-09-648-310-4 (1-81) x US-09-938-842A-355 (1-486)

```
    1 MetAsnValAspHisGluValAsnLeuLeuValGluGluIleHisArgLeuGlySerLys 20
      ||||::||||||||| |||  ||  ::::|||||||||||||||||||||::::
  199 ATGAACGTAGACGAAGAGAGATTCAGAAACTGGAAGAAGAGATCCATCGTCTTGGTTCTCGT 258

   21 AsnAlaAspGlyLysLeuSerValLysPheGlyValLeuPheArgAspLysCysAla 40
      |||||||||||||||||||||||||||:|||  ::::|||||
  259 CAGACCGATGGCTCTTACAAGGTGACGTTTGGAGTGTTCAATGACGATCGATGTGCC 318

   41 AsnLeuPheGluAlaLeuValGlyThrLeuLysAlaAlaLysArgArgLysIleValThr 60
      ||||::||||||||::::||||||||||||||::::|||::::
  319 AACATCTTTGAAGCATTGGTTGGGACTCTGAGAGCTGCCAAGAAACGCAAAATAGTCGCA 378

   61 TyrProGlyGluLeuLeuLeuGlnGlyValHisAspValAspIleIleLeu 78
      ::: ||||||||||||||||||||||||||||
  379 TTTGAAGGTGAACTTTGCTTCAAGGCGTTCACGATAAGGTTGAGATCACTCTC 432
```

SULT 10
-09-770-961-618/c
Sequence 618, Application US/09770961
Publication No. US20030115639A1
GENERAL INFORMATION:
, APPLICANT: Gorlach, Jorn
, APPLICANT: An, Yong-Qiang
, APPLICANT: Hamilton, Carol M.
, APPLICANT: Price, Jennifer L.
, APPLICANT: Raines, Tracy M.
, APPLICANT: Yu, Yang
, APPLICANT: Rameaka, Joshua G.
, APPLICANT: Page, Amy
, APPLICANT: Matthew, Abraham V.
, APPLICANT: Ledford, Brooke L.
, APPLICANT: Woessner, Jeffrey P.
, APPLICANT: Haas, William David
, APPLICANT: Garcia, Carlos A.
, APPLICANT: Kricker, Maja

, APPLICANT: Slader, Ted
, APPLICANT: Davis, Keith R.
, APPLICANT: Allen, Keith
, APPLICANT: Hoffman, Neil
, APPLICANT: Hurban, Patrick
, TITLE OF INVENTION: Expressed Sequences of Arabidopsis
, TITLE OF INVENTION: thaliana
, FILE REFERENCE: 2026 (PARA-015PRV)
, CURRENT APPLICATION NUMBER: US/09/770,961
, CURRENT FILING DATE: 2001-01-26
, PRIOR APPLICATION NUMBER: 60/178,466
, PRIOR FILING DATE: 2000-01-27
, NUMBER OF SEQ ID NOS: 999
, SOFTWARE: FastSEQ for Windows Version 4.0
, SEQ ID NO 618
,   LENGTH: 492
,   TYPE: DNA
,   ORGANISM: Arabidopsis thaliana
US-09-770-961-618

Alignment Scores:
Pred. No.:   1.49e-37   Length: 492
Score: 286.00   Matches: 55
Percent Similarity: 80.77%   Conservative: 8
Best Local Similarity: 70.51%   Mismatches: 15
Query Match: 69.76%   Indels: 0
DB: 10   Gaps: 0

US-09-648-310-4 (1-81) x US-09-770-961-618 (1-492)

```
Qy    1 MetAsnValAspHisGluValAsnLeuLeuValGluGluIleHisArgLeuGlySerLys 20
        ||||::||||||||| |||  ||  ::::|||||||||||||||||::::
Db  457 ATGAACGTAGACGAAGAGATTCAGAAACTGGAAGAAGAGATCCATCGTCTTGGTTCTCGT 398

Qy   21 AsnAlaAspGlyLysLeuSerValLysPheGlyValLeuPheArgAspLysCysAla 40
        ||||||||||||||||||||||||||:|||  ::::|||||
Db  397 CAGACCGATGGCTCTTACAAGGTGACGTTTGGAGTGTTCAATGACGATCGATGTGCC 338

Qy   41 AsnLeuPheGluAlaLeuValGlyThrLeuLysAlaAlaLysArgArgLysIleValThr 60
        ||||::||||||::::||||||||||||::::|||::::
Db  337 AACATCTTTGAAGCATTGGTTGGGACTCTGAGAGCTGCCAAGAAACGCAAAATAGTCGCA 278

Qy   61 TyrProGlyGluLeuLeuLeuGlnGlyValHisAspValAspIleIleLeu 78
        ::: ||||||||||||||||||||||||||||
Db  277 TTTGAAGGTGAACTTTGCTTCAAGGCGTTCACGATAAGGTTGAGATCACTCTC 224
```

RESULT 11
US-10-424-599-94878/c
; Sequence 94878, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
, APPLICANT: La Rosa Thomas J
, APPLICANT: Kovalic David K
, APPLICANT: Zhou Yihua
, APPLICANT: Cao Yongwei
, TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
, TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
, FILE REFERENCE: 38-21(53223)B
, CURRENT APPLICATION NUMBER: US/10/424,599
, CURRENT FILING DATE: 2003-04-28
, NUMBER OF SEQ ID NOS: 285684
, SEQ ID NO 94878
,   LENGTH: 716
,   TYPE: DNA
,   ORGANISM: Glycine max
,   FEATURE:
,   NAME/KEY: unsure
,   LOCATION: (1)..(716)
,   OTHER INFORMATION: unsure at all n locations
,   FEATURE:
,   OTHER INFORMATION: Clone ID: PAT_MRT3847_56687C.1
US-10-424-599-94878

Alignment Scores:
Pred. No.: 1.73e-34   Length: 716
Score: 269.00   Matches: 52
Percent Similarity: 79.49%   Conservative: 10
Best Local Similarity: 66.67%   Mismatches: 16
Query Match: 65.61%   Indels: 0
DB: 12   Gaps: 0

US-09-648-310-4 (1-81) x US-10-424-599-94878 (1-716)

Qy   1 MetAsnValAspHisGluValAsnLeuLeuValGluGluIleHisArgLeuGlySerLys 20
            |||||:||:| ::|||||||| |||||| |||||||
Db 537 ATGAATGTTGAAGAAGAGGTTCAGCGTCTCAGCCAAGAAATCAAGAGGCTTGGTAAGGTC 478

Qy   21 AsnAlaAspGlyLysLeuSerValLysPheGlyValLeuPheArgAspAspLysCysAla 40
            |||||||||||| ||||| |||:   ::|||||
Db 477 CAAGCAGATGGTTCTTACAAGGTAACATTTGGAACACTATTTAACGATGAATGCGCA 418

Qy   41 AsnLeuPheGluAlaLeuValGlyThrLeuLysArgAlaAlaLysArgLysIleValThr 60
            |||| :|||||| |||||||| |:|||:::|||| :::||||| :::|||
Db 417 AATATATTTGAAGCACTTGTTGGAACACTAAGAGCAGCGAAAAAGCGTAAAGTACTGACA 358

Qy   61 TyrProGlyLeuLeuLeuGlnGlyValHisAspAspValAspIleIleLeu 78
            ||||||||||||| :::||| ||| ||:|:|||||||
Db 357 TACGGAGGTGAGCTACTGCTCCAAGGAGTCCATGATAATGTGGAAATCACTCTT 304

RESULT 13
US-10-033-528-939/c
Sequence 939, Application US/10033528
GENERAL INFORMATION:
; Publication No. US20020131971A1
; GENERAL INFORMATION:
; APPLICANT: King, Gordon E.
; APPLICANT: Meagher, Madeleine Joy
; APPLICANT: Xu, Jiangchun
; APPLICANT: Secrist, Heather
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF COLON CANCER
; FILE REFERENCE: 210121.547C1
; CURRENT APPLICATION NUMBER: US/10/033,528
; CURRENT FILING DATE: 2001-12-26
; NUMBER OF SEQ ID NOS: 1896
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 939
; LENGTH: 552
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-033-528-939

Alignment Scores:
Pred. No.: 7.79e-32   Length: 552
Score: 252.00   Matches: 50
Percent Similarity: 100.00%   Conservative: 0
Best Local Similarity: 100.00%   Mismatches: 0
Query Match: 61.46%   Indels: 0
DB: 13   Gaps: 0

US-10-033-528-939 (1-81) x US-10-033-528-939 (1-552)

Qy   32 ValLeuPheArgAspAspLysCysAlaAsnLeuPheGluAlaLeuValGlyThrLeuLys 51
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 552 GTCCTCTTCCGTGATGATAAATGTGCCAACCTCTTTGAAGCATTGGTAGGAACTCTTAAA 493

Qy   52 AlaAlaLysArgArgLysIleValThrTyrProGlyLeuLeuLeuGlnGlyValHis 71
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 492 GCTGCAAAACGAAGGAAGATTGTAACATATCCAGGAGAGCTGCTTCTCGCAAGGTGTTCAT 433

Qy   72 AspAspValAspIleIleLeuLeuGlnAsp 81
            ||||||||||||||||||||||||
Db 432 GATGATGTTGACATTATATTACTGCAAGAT 403

RESULT 14
US-10-099-926-939/c
; Sequence 939, Application US/10099926
; Publication No. US20030166064A1
; GENERAL INFORMATION:
; APPLICANT: King, Gordon E.
; APPLICANT: Meagher, Madeleine Joy
; APPLICANT: Xu, Jiangchun
; APPLICANT: Secrist, Heather
; APPLICANT: Jiang, Yuqiu
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF COLON CANCER
; FILE REFERENCE: 210121.547C2
; CURRENT APPLICATION NUMBER: US/10/099,926
; CURRENT FILING DATE: 2002-03-17
; NUMBER OF SEQ ID NOS: 1982
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 939
; LENGTH: 552
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-099-926-939

Alignment Scores:
Pred. No.: 7.79e-32   Length: 552
Score: 252.00   Matches: 50
Percent Similarity: 100.00%   Conservative: 0
Best Local Similarity: 100.00%   Mismatches: 0
Query Match: 61.46%   Indels: 0
DB: 14   Gaps: 0

US-09-648-310-4 (1-81) x US-10-099-926-939 (1-552)

Qy   32 ValLeuPheArgAspAspLysCysAlaAsnLeuPheGluAlaLeuValGlyThrLeuLys 51
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 552 GTCCTCTTCCGTGATGATAAATGTGCCAACCTCTTTGAAGCATTGGTAGGAACTCTTAAA 493

Qy   52 AlaAlaLysArgArgLysIleValThrTyrProGlyLeuLeuLeuGlnGlyValHis 71
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 492 GCTGCAAAACGAAGGAAGATTGTAACATATCCAGGAGAGCTGCTTCTCGCAAGGTGTTCAT 433

Qy   72 AspAspValAspIleIleLeuLeuGlnAsp 81
            ||||||||||||||||||||||||
Db 432 GATGATGTTGACATTATATTACTGCAAGAT 403

RESULT 12
US-09-920-300A-939/c
Sequence 939, Application US/09920300A
Patent No. US2002013 6728A1
GENERAL INFORMATION:
APPLICANT: King, Gordon E.
APPLICANT: Meagher, Madeleine Joy
APPLICANT: Xu, Jiangchun
APPLICANT: Secrist, Heather
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
TITLE OF INVENTION: AND DIAGNOSIS OF COLON CANCER
FILE REFERENCE: 210121.547
CURRENT APPLICATION NUMBER: US/09/920,300A
CURRENT FILING DATE: 2001-07-31
NUMBER OF SEQ ID NOS: 1789
SOFTWARE: FastSEQ for Windows Version 4.0
SEQ ID NO 939
LENGTH: 552
TYPE: DNA
ORGANISM: Homo sapiens
US-09-920-300A-939

Alignment Scores:
Pred. No.: 7.79e-32   Length: 552
Score: 252.00   Matches: 50
Percent Similarity: 100.00%   Conservative: 0
Best Local Similarity: 100.00%   Mismatches: 0
Query Match: 61.46%   Indels: 0
DB: 9   Gaps: 0

US-09-648-310-4 (1-81) x US-09-920-300A-939 (1-552)

Qy   32 ValLeuPheArgAspAspLysCysAlaAsnLeuPheGluAlaLeuValGlyThrLeuLys 51
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 552 GTCCTCTTCCGTGATGATAAATGTGCCAACCTCTTTGAAGCATTGGTAGGAACTCTTAAA 493

Qy   52 AlaAlaLysArgArgLysIleValThrTyrProGlyLeuLeuLeuGlnGlyValHis 71
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 492 GCTGCAAAACGAAGGAAGATTGTAACATATCCAGGAGAGCTGCTTCTCGCAAGGTGTTCAT 433

Qy   72 AspAspValAspIleIleLeuLeuGlnAsp 81
            ||||||||||||||||||||||||
Db 432 GATGATGTTGACATTATATTACTGCAAGAT 403

ESULT 13
S-10-033-528-939/c
Sequence 939, Application US/10033528

```
 32 ValLeuPheArgAspAspLysCysAlaAsnLeuPheGluAlaLeuValGlyThrLeuLys 51
    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
552 GTCCTCTTCCGTGATGATAAAATGTGCCAACCTCTTGAAGCATTGGTAGGAACTCTTAAA 493

 52 AlaAlaLysArgArgLysIleValThrTyrProGlyGluLeuLeuGlnGlyValHis 71
    |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
492 GCTGCAAAACGAAGAAGATTGTAACATATCCAGGAGAGCTGCTTCTGCAAGGTGTTCAT 433

 72 AspAspValAspIleIleLeuLeuGlnAsp 81
    ||||||||||||||||||||||||||||||
432 GATGATGTTGACATTATTACTGCAAGAT 403
```

SULT 15
-09-294-093B-2020
Sequence 2020, Application US/09294093B
Patent No. US20010051335A1
GENERAL INFORMATION:
APPLICANT: Lalgudi, Raghunath, V.
APPLICANT: Ito, Laura, Y.
APPLICANT: Sherman, Bradley, K.
TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN TASSEL
FILE REFERENCE: PL-0009 US
CURRENT APPLICATION NUMBER: US/09/294,093B
CURRENT FILING DATE: 1999-04-16
PRIOR APPLICATION NUMBER: 60/082,567
PRIOR FILING DATE: April 21, 1998
NUMBER OF SEQ ID NOS: 6207
SOFTWARE: PERL Program
SEQ ID NO 2020
  LENGTH: 268
  TYPE: DNA
  ORGANISM: Zea mays
  FEATURE:
  NAME/KEY: misc_feature
  OTHER INFORMATION: Incyte ID No. US20010051335A1 700345395H1
-09-294-093B-2020

Alignment Scores:
Seed. No.:   8.17e-29   Length: 268
Score: 231.00   Matches: 44
Percent Similarity: 80.95%   Conservative: 7
Best Local Similarity: 69.84%   Mismatches: 12
Query Match: 56.34%   Indels: 0
:   9   Gaps: 0

-09-648-310-4 (1-81) x US-09-294-093B-2020 (1-268)

```
  1 MetAsnValAspHisGlyValAsnLeuValGluGluIleHisArgLeuGlySerLys 20
    |||||||||:::   |||||||||||||||||||||||||||||||||   :::
 79 ATGAACGTGGAGGAGGAGTCGGGAGGCTCAAGGAGGAGATCCATAGGCTCGGCCAGCAG 138

 21 AsnAlaAspGlyLysLeuSerValLysPheGlyValLeuPheArgAspAspLysCysAla 40
    |||||||||||||||||||||||||||||||::::|||||||||||||||::::||||||
139 CAGCCCGATGGCTCTTCAAGGTCAAGTTGGCGTTCTTCAACGACGATCGGTGTGCA 198

 41 AsnLeuPheGluAlaLeuValGlyThrLeuLysAlaAlaLysArgArgLysIleValThr 60
    |||::::|||||||||||||||||||::|||||||||||||||||::::|||::::|
199 AATATCTTTGAAGCACTAGTTGGCACCTTGAGGGCCGCCAAGAGAGGAAGATCTTGACC 258

 61 TyrProGly 63
    |||||||||
259 TACGAAGGC 267
```

Search completed: March 17, 2004, 17:51:12
Job time : 1242 secs

```
                GenCore version 5.1.6
        Copyright (c) 1993 - 2004 Compugen Ltd.

protein - nucleic search, using frame_plus_p2n model

Run on:   March 17, 2004, 12:34:46 ; Search time 318 Seconds
                              (without alignments)
                              1082.088 Million cell updates/sec

title:         US-09-648-310-4
Perfect score: 410
Sequence:      1 MNVDHEVNLLVBEIHRLGSK........PGELLLQGVHDDVDIILLQD 81

Scoring table: BLOSUM62
               Xgapop 10.0 , Xgapext 0.5
               Ygapop 10.0 , Ygapext 0.5
               Fgapop  6.0 , Fgapext 7.0
               Delop   6.0 , Delext  7.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters:    6747726

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Command line parameters:
CODEL=frame+_p2n.model -DEV=xlh
=/cgn2_1/USPTO_spool/US09648310/runat_15032004_101253_16551/app_query.fasta_1.263
B=N Geneseq_29Jan04 -QFMT=fastap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPCL=0
COPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
IST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
CODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
SER=US09648310_@CGN_1_1470 @runat_15032004_101253_16551 -NCPU=6 -ICPU=3
O_MMAP -LARGEQUERY -NEG_SCORES=3 -WAIT -DSPBLOCK=100 -LONGLOG
EV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
GAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7
```

```
Database: N_Geneseq_29Jan04:*
   1: geneseqm1980s:*
   2: geneseqm1990s:*
   3: geneseqm2000s:*
   4: geneseqm2001as:*
   5: geneseqm2001bs:*
   6: geneseqm2002s:*
   7: geneseqm2003as:*
   8: geneseqm2003bs:*
   9: geneseqm2003cs:*
  10: geneseqm2004s:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.
```

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	410	100.0	627	6	ABQ56862	Abq56862 Human col
2	410	100.0	712	3	AAC66413	Aac66413 Human sec
3	410	100.0	778	3	AAF16110	Aaf16110 Human pro
4	410	100.0	835	6	ABK11086	Abk11086 Human Pro
5	406	99.0	594	6	ABQ57055	Abq57055 Human col
6	404	98.5	717	6	ABQ57053	Abq57053 Human col
7	390	95.1	780	6	ABK11085	Abk11085 Rat Progr
8	387	94.4	458	8	ACH38863	Ach38863 Human foe

No.	Score	bits	Length	DB	ID	Description
9 c	380	92.7	653	6	ABT10814	Human bre
10	352	85.9	800	2	AAZ21517	Progressi
11	295	72.0	551	3	AAC53910	Zea mays
12 c	286	69.8	486	9	ABZ12550	Arabidops
13	286	69.8	492	3	ADE81847	Arabidops
14	286	69.8	657	9	AAC53673	Arabidops
15 c	252	61.5	552	6	ABK45388	cDNA enco
16	231	56.3	268	6	ABL72646	Corn tass
17 c	216	52.7	388	3	AAC98308	Human col
18	214.5	52.3	351	2	ABA11066	Human ner
19	149	36.3	289	6	ABL74776	Corn tass
20	105	25.6	60	6	ABN49354	Human spl
21	102	24.9	1322	9	ADB63045	Human cDN
22	102	24.9	1373	5	AAS92182	DNA encod
23	86	21.0	175	6	ABL76064	Corn tass
24	78.5	19.1	552	4	ABL04119	Drosophil
25 c	78.5	19.1	2610	4	ABL04118	Drosophil
26	74.5	18.2	90600	6	ABQ78872	S. roseos
27	74	18.0	1649	6	ABL06465	Drosophil
28	74	18.0	7090	4	ABL06464	Drosophil
29 c	73.5	17.9	2323	7	AAQ57966	Bacterial
30 c	71.5	17.4	91608	7	AAL54538	Arabidops
31 c	70	17.1	1728	7	ACA28152	Prokaryot
32	66	16.1	481	7	ACD94535	Human col
33	66	16.1	1084	6	ABQ42188	Oligonucl
34 c	66	16.1	1084	6	ABQ42189	Oligonucl
35 c	66	16.1	1257	7	ACA27028	Prokaryot
36	66	16.1	2133	8	ADA30487	DNA encod
37	66	16.1	201143	6	ABK83568	Human DNA
38	64.5	15.7	410	6	ABV32849	Human pro
39	64.5	15.7	1224	7	ACA57215	Human adi
40	64.5	15.7	1901	3	AAA47405	1-decoxy-D
41 c	64.5	15.7	2323	2	AAQ76290	Chitinase
42 c	64.5	15.7	2323	2	AAT49479	Serratia
43 c	64.5	15.7	2323	2	AAT86836	DNA encod
44 c	64.5	15.7	2323	2	AAV35601	Serratia
45 c	64.5	15.7	2331	3	AAA08583	Human cyt

ALIGNMENTS

```
RESULT 1
ABQ56862/c
ID   ABQ56862 standard; cDNA; 627 BP.
XX
AC   ABQ56862;
XX
DT   02-AUG-2002  (first entry)
XX
DE   Human colon cancer related nucleotide sequence SEQ ID NO:557.
XX
KW   Human; colon cancer; cancer; tissue profiling; forensic; mapping;
KW   genetic analysis; diagnostic; antisense therapy; gene; ss.
XX
OS   Homo sapiens.
XX
PN   WO200229086-A2.
XX
PD   11-APR-2002.
XX
PF   02-OCT-2001; 2001WO-US030732.
XX
PR   02-OCT-2000; 2000US-0237271P.
XX
PA   (FARB ) BAYER CORP.
XX
PI   Burgess C, Astle JH, Carroll E, Catino TJ, Dwivedi P, Molino GA;
PI   Thiaglingam A, Lewis ME;
XX
DR   WPI; 2002-426115/45.
XX
PT   New isolated nucleic acid that is differentially expressed in cancer
PT   tissues useful for determining the presence of colon cancer in a cell or
```

Not 102(e)

tissue type, and in antisense therapy.

Claim 1; Fig 1; 796pp; English.

ABQ56306 to ABQ60787 represent isolated nucleic acids (I) differentially expressed in cancer tissues. ABB78993 to ABB79004 represent proteins encoded by the ABQ60776 to ABQ60787 nucleic acid sequences. (I) can be used in antisense therapy. An antibody immunoreactive with a polypeptide encoded by (I) is useful for detecting cancer in a patient sample, and for detecting the presence or absence of a polynucleotide encoded by a nucleic acid which hybridises to (I) in a cell. A probe/primer derived from (I) can be used for determining the presence of a nucleic acid which hybridises to (I), and for determining the phenotype of cells in a sample of cells from a patient. (I) is useful for determining the presence of colon cancer in a cell or tissue type, for determining the presence or state of other type of cancer, in antisense therapy, to generate macroarrays on a solid surface, to identify a chromosome on which the corresponding gene resides, and in tissue profiling, forensics, genetic analysis, mapping and diagnostic application. (I) can be used to raise antibodies, and to screen for peptide analogues and antagonists

Sequence 627 BP; 171 A; 151 C; 153 G; 147 T; 0 U; 5 Other;

Alignment Scores:
```
Pred. No.:                        Length:           627
Score:                   2.29e-53 Matches:           81
Percent Similarity:      410.00   Conservative:       0
Best Local Similarity:   100.00%  Mismatches:         0
Query Match:             100.00%  Indels:             0
DB:                               Gaps:               6
```

US-09-648-310-4 (1-81) x ABQ56862 (1-627)

```
Qy    1 MetAsnValAspHisGluValAsnLeuLeuValGluGluIleHisArgLeuGlySerLys 20
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  271 ATGAATGTGGATCACGAGGTTAACCTCTTAGTGGAGGAAATTCATCGTTTGGGTTCAAAA 212

Qy   21 AsnAlaAspGlyLysLeuSerValLysPheGlyValLeuPheArgAspAspLysCysAla 40
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  211 AATGCTGATGAAAGTTAAGCGTGAAATTTGGGGTCCTCTTCCGTGATGATAAATGTGCC 152

Qy   41 AsnLeuPheGluAlaLeuValGlyThrLeuLysAlaAlaLysArgArgLysIleValThr 60
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  151 AACCTCTTTGAAGCATTGGTAGGAACTCTTAAAGCTGCAAAACGAAGAAGATTGTAACA 92

Qy   61 TyrProGlyGluLeuLeuGlnGlyValHisAspAspValAspIleIleLeuLeuGln 80
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   91 TATCCAGGAGAGCTGCTTCTGCAAGGTGTTCATGATGATGTTGACATTATATTACTGCAA 32

Qy   81 Asp 81
          |||
Db   31 GAT 29
```

RESULT 2

AC66413

ID AAC66413 standard; DNA; 712 BP.
XX
AC AAC66413;
XX
DT 14-FEB-2001 (first entry)
XX
DE Human secreted protein coding sequence SEQ ID NO: 14.
XX
KW Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
KW antiallergic; hepatotropic; antidiabetic; antiinflammatory; antiulcer;
KW vulnerary; anticonvulsant; antibacterial; antifungal; antiparasitic;
KW cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;
KW neurological disease; infection; human; secreted protein; ss.
XX
OS Homo sapiens.
XX
XX
    WO200058350-A1.

PD 05-OCT-2000.
XX
PF
PR 22-MAR-2000; 2000WO-US007483.
XX
PR 26-MAR-1999; 99US-0126596P.
PR 22-DEC-1999; 99US-0171552P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Ruben SM, Komatsoulis G;
XX
DR WPI; 2000-602357/57.
DR P-PSDB; AAB32005.
XX
PT Nucleic acid molecules encoding human secreted proteins, used in
PT preventing, treating or ameliorating a disorder, e.g. Alzheimer's and
PT Parkinson's diseases and cancers.
XX
PS Claim 1; Page 336; 423pp; English.
XX
CC The invention relates to the isolation of genes AAC66410-C66458 encoding
CC 49 human secreted proteins AAB32002-B32050. The genes can be used to
CC generate fusion proteins by linking to the gene for the human
CC immunoglobulin G Fc portion (SEQID1) for increasing the stability of the
CC fusion protein as compared to the human protein only. The genes and
CC proteins are useful for preventing, ameliorating or treating medical
CC conditions, e.g. by protein or gene therapy. The genes are isolated from
CC a range of human tissues disclosed in the specification. The nucleic
CC acids, proteins, antibodies and (ant)agonists are useful in the
CC diagnosis, treatment and prevention of: (a) cancer, e.g. breast and
CC ovarian cancer, and other cancers of the adrenal gland, bone, bone
CC marrow, breast, gastrointestinal tract, liver, lung, or urogenital; (b)
CC immune disorders e.g. Addison's disease, allergies, autoimmune haemolytic
CC anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,
CC multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c)
CC cardiovascular disorders such as myocardial ischaemias; (d) wound healing
CC ; (e) neurological diseases e.g. cerebral anoxia and epilepsy; and (f)
CC infectious diseases such as viral, bacterial, fungal and parasitic
CC infections
XX
SQ Sequence 712 BP; 241 A; 121 C; 149 G; 201 T; 0 U; 0 Other;

Alignment Scores:
```
Pred. No.:                        Length:           712
Score:                   2.74e-53 Matches:           81
Percent Similarity:      410.00   Conservative:       0
Best Local Similarity:   100.00%  Mismatches:         0
Query Match:             100.00%  Indels:             0
DB:                               Gaps:               3
```

US-09-648-310-4 (1-81) x AAC66413 (1-712)

```
Qy    1 MetAsnValAspHisGluValAsnLeuLeuValGluGluIleHisArgLeuGlySerLys 20
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  129 ATGAATGTGGATCACGAGGTTAACCTCTTAGTGGAGGAAATTCATCGTTTGGGTTCAAAA 188

Qy   21 AsnAlaAspGlyLysLeuSerValLysPheGlyValLeuPheArgAspAspLysCysAla 40
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  189 AATGCTGATGGAAAGTTAAGCGTGAAATTTGGGGTCCTCTTCCGTGATGATAAATGTGCC 248

Qy   41 AsnLeuPheGluAlaLeuValGlyThrLeuLysAlaAlaLysArgLysIleValThr 60
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  249 AACCTCTTTGAAGCATTGGTAGGAACTCTTAAAGCTGCAAAACGAAGAAGATTGTAACA 308

Qy   61 TyrProGlyGluLeuLeuGlnGlyValHisAspAspValAspIleIleLeuLeuGln 80
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  309 TATCCAGGAGAGCTGCTTCTGCAAGGTGTTCATGATGATGTTGACATTATATTACTGCAA 368

Qy   81 Asp 81
          |||
Db  369 GAT 371
```

RESULT 3

```
Db    246  AATGCTGATGAAAGTTAAGGCTGAAATTTGGGGTCCTTCCGTGATGATAAATGTGCC  305
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy     41  AsnLeuPheGluAlaLeuValGlyThrLeuLysAlaAlaLysArgArgLysIleValThr  60

Db    306  AACCTCTTGAAGCATTGGTAGGAACTCTTAAAGCTGCAAAACGAGGAAGATTGTAACA  365
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy     61  TyrProGlyGluLeuLeuLeuGlnGlyValHisAspAspValAspIleIleLeuLeuGln  80

Db    366  TATCCAGGAGAGCTGCTTCTGCAAGGTGTTCATGATGATGTTGACATTATTATTACTGCAA  425
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy     81  Asp 81
                |||
Db    426  GAT 428
```

```
RESULT 4
ABK11086
ID   ABK11086 standard; cDNA; 835 BP.
XX
AC   ABK11086;
XX
DT   05-JUN-2002  (first entry)
XX
DE   Human Progression Suppressed Gene 13 (HuPSGen 13), cDNA.
XX
KW   Human; Progression Suppressed Gene 13; rPSGen 13; cytostatic; cancer;
KW   blood vessel; nasopharyngeal tumour; thyroid tumour; leukaemia; lymphoma;
KW   breast; lung; prostate; ovary; colon; gene; ss.
XX
OS   Homo sapiens.
XX
FH   Key            Location/Qualifiers
FT   CDS            197..442
FT                  /*tag= a
FT                  /product= "Progression suppressed gene 13 protein"
XX
PN   WO200216419-A2.
XX
PD   28-FEB-2002.
XX
PF   27-AUG-2001; 2001WO-US026795.
XX
PR   25-AUG-2000; 2000US-00648310.
XX
PA   (UYCO ) UNIV COLUMBIA NEW YORK.
XX
PI   Fisher PB, Kang D,  Su Z;
XX
DR   WPI; 2002-280914/32.
DR   P-PSDB; AAU76533.
XX
PT   New rat and human Progression Suppressed Gene 13 for preventing the
PT   growth of cancer cells and/or new blood vessels, and for treating
PT   patients suffering from a cancer.
XX
PS   Claim 4; Fig 2; 53pp; English.
XX
CC   The invention relates to novel isolated nucleic acids which encode a rat
CC   or human Progression Suppressed Gene 13 (PSGen 13) protein. The nucleic
CC   acids are useful for preventing the growth of cancer cells and/or new
CC   blood vessels, and for treating patients suffering from a cancer, e.g.
CC   nasopharyngeal tumour, thyroid tumour, leukaemia, lymphoma, or cancer of
CC   the breast, lung, prostate, ovary or colon. PSGen 13 may also be used to
CC   suppress the transformed phenotype of a malignant cell. Administration of
CC   PSGen 13 gene or protein may result in a decrease in tumour mass, number
CC   of cancer cells, serum tumour marker, tumour metastasis, vascularisation,
CC   perfusion, or rate of tumour growth, improved clinical symptoms, and/or
CC   increased patient survival. The present sequence represents the coding
CC   sequence of human Progression Suppressed Gene 13 (HuPSGen 13)
XX
SQ   Sequence 835 BP; 246 A; 160 C; 176 G; 253 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                                          835
```

```
16110
AAF16110 standard; cDNA; 778 BP.

AAF16110;

13-MAR-2001  (first entry)

Human prostate cancer antigen nucleotide sequence SEQ ID NO:545.

Human; prostate cancer; prostate cancer antigen; detection; diagnosis;
neuroprotective; cytostatic; cardioactive; immunomodulatory; muscular;
vulnerary; gastrointestinal; nephrotropic; antiinfective; gynaecological;
antibacterial; gene therapy; neural; immune; reproductive; renal;
gastrointestinal; pulmonary; cardiovascular; proliferative disorder;
wound; infectious disease; ss.

Homo sapiens.

WO20005517A-A1.

21-SEP-2000.

08-MAR-2000; 2000WO-US005988.

12-MAR-1999;  99US-0124270P.

(HUMA-) HUMAN GENOME SCI INC.
(ROSE/) ROSEN C A.

Rosen CA,  Ruben SM;

WPI; 2000-587513/55.
P-PSDB; AAB56907.

Prostate cancer associated gene sequences, referred to as prostate cancer
antigens, useful for treatment, prevention, and diagnosis of disorders
such as prostate cancer.

Claim 1; Page 1004-1005; 2338pp; English.

AAF15566 to AAF16505 encode the human prostate cancer associated
proteins, called prostate cancer antigens, given in AAB56363 to AAB57302.
The prostate cancer antigens can have neuroprotective, cytostatic,
cardioactive, immunomodulatory, muscular, vulnerary, gastrointestinal,
nephrotropic, antiinfective, gynaecological and antibacterial activities,
and can be used in gene therapy. The prostate cancer antigen
polynucleotides may be used for detection of prostate cancer, chromosome
identification, as chromosome markers, and for numerous other diagnostic
or research purposes. The prostate cancer antigens may be used to treat
disorders such as neural, immune, muscular, reproductive, renal,
gastrointestinal, pulmonary, cardiovascular, renal, and proliferative
disorders, wounds, and infectious diseases. AAF16506 to AAF16514 to
AAB57303 represent sequences used in the exemplification of the present
invention
```

```
Sequence 778 BP; 250 A; 151 C; 148 G; 221 T; 0 U; 8 Other;

Alignment Scores:
Pred. No.:                     Length:          778
Score:            3.11e-53     Matches:          81
                  410.00       Conservative:      0
Percent Similarity:  100.00%   Mismatches:        0
Percent Local Similarity: 100.00%  Indels:        0
Query Match:      100.00%      Gaps:              3
                  3
```

```
-09-648-310-4 (1-81) x AAF16110 (1-778)

  1  MetAsnValAspHisGluValAsnLeuValGluGluIleHisArgLeuGlySerLys  20
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||
186  ATGAATGTGGATCACGAGGTTAACCTCTTAGTGGAGGAAATTCATCGTTTGGGTTCAAAA  245

 21  AsnAlaAspGlyLysLeuSerValLysPheGlyValLeuPheArgAspAspLysCysCysAla  40
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

Score: 410.00   Matches: 81
Percent Similarity: 100.00%   Conservative: 0
Best Local Similarity: 100.00%   Mismatches: 0
Query Match: 100.00%   Indels: 0
DB: Gaps: 6

us-09-648-310-4 (1-81) x ABK11086 (1-835)

```
Qy    1 MetAsnValAspHisGluValAsnLeuLeuValGluGluIleHisArgLeuGlySerLys  20
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  197 ATGAATGTGGATCACGAGGTTAACCTCTTAGTGGAGGAAATTCATCGTTTGGGTTCAAAA 256

Qy   21 AsnAlaAspGlyLysLeuSerValLysPheGlyValLeuPheArgAspAspLysCysAla  40
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  257 AATGCTGATGGAAAGTTAAGCGTGAAAATTTGGGGTCCTCTTCCGTGATGATAAATGTGCC 316

Qy   41 AsnLeuPheGluAlaLeuValGlyThrLeuLysAlaAlaLysArgArgLysIleValThr  60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  317 AACCTCTTTGAAGCATTGGTAGGAACTCTTAAAGCTGCAAAACGAAGGAAGATTGTAACA 376

Qy   61 TyrProGlyGluLeuLeuGlnGlyValHisAspAspValAspIleIleLeuLeuLeuGln  80
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  377 TATCCAGGAGAGCTGCTTCTGCAAGGTGTTCATGATGATGTTGACATTATTATTACTGCAA 436

Qy   81 Asp  81
          |||
Db  437 GAT 439
```

RESULT 5
ABQ57055
ID   ABQ57055 standard; cDNA; 594 BP.
XX
AC   ABQ57055;
XX
DT   02-AUG-2002 (first entry)
XX
DE   Human colon cancer related nucleotide sequence SEQ ID NO:750.
XX
KW   Human; colon cancer; cancer; tissue profiling; forensic; mapping;
KW   genetic analysis; diagnostic; antisense therapy; gene; ss.
XX
OS   Homo sapiens.
XX
PN   WO20229086-A2.
XX
PD   11-APR-2002.
XX
PF   02-OCT-2001; 2001WO-US030732.
XX
PR   02-OCT-2000; 2000US-0237271P.
XX
PA   (FARB) BAYER CORP.
XX
PI   Burgess C, Astle JH, Carroll E, Catino TJ, Dwivedi P, Molino GA;
PI   Thiaglingam A, Lewis ME;
XX
DR   WPI; 2002-426115/45.
XX
CS   Claim 1; Fig 1; 796pp; English.
XX
CC   ABQ56306 to ABQ60787 represent isolated nucleic acids (I) differentially
CC   expressed in cancer tissues. ABB78993 to ABB79004 represent proteins
CC   encoded by the ABQ60076 to ABQ60787 nucleic acid sequences. (I) can be
CC   used in antisense therapy. An antibody immunoreactive with a polypeptide
CC   encoded by (I) is useful for detecting cancer in a patient sample, and
CC   for detecting the presence or absence of a polynucleotide encoded by a
CC   nucleic acid which hybridises to (I) in a cell. A probe/primer derived
CC   from (I) can be used for determining the presence of a nucleic acid which
CC   hybridises to (I), and for determining the phenotype of cells in a sample

---

CC   of cells from a patient. (I) is useful for determining the presence of
CC   colon cancer in a cell or tissue type, for determining the presence or
CC   state of other type of cancer, in antisense therapy, to generate
CC   macroarrays on a solid surface, to identify a chromosome on which the
CC   corresponding gene resides, and in tissue profiling, forensics, genetic
CC   analysis, mapping and diagnostic applications. (I) can be used to raise
CC   antibodies, and to screen for peptide analogues and antagonists
XX
SQ   Sequence 594 BP; 157 A; 142 C; 134 G; 149 T; 0 U; 12 Other;

Alignment Scores:
Pred. No.:
Score: 8.8e-53   Length: 594
Percent Similarity: 406.00   Matches: 80
Best Local Similarity: 100.00%   Conservative: 1
Query Match: 98.77%   Mismatches: 0
DB: 99.02%   Indels: 0
                                         Gaps: 0

US-09-648-310-4 (1-81) x ABQ57055 (1-594)

```
Qy    1 MetAsnValAspHisGluValAsnLeuLeuValGluGluIleHisArgLeuGlySerLys  20
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  199 ATGAATGTGGATCACGAGGTTAACCTCTTAGTGGAGGAAATTCATCGTTTGGGTTCAAAA 258

Qy   21 AsnAlaAspGlyLyLeuSerValLysPheGlyValLeuPheArgAspAspLysCysAla  40
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  259 AATGCTGATGGAAAGTTAAGCGTGAAAATTTGGGGTCCTCTTCCGTGATGATAAATGTGCC 318

Qy   41 AsnLeuPheGluAlaLeuValGlyThrLeuLysAlaAlaLysArgArgLysIleValThr  60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  319 AACCTCTTTGAAGCATTGGTAGGAACTCTTAAAGCTGCAAAACGAAGGAAGATTGTAACA 378

Qy   61 TyrProGlyGluLeuLeuGlnGlyValHisAspAspValAspIleIleLeuLeuLeuGln  80
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  379 TATCCAGGAGAGCTGCTTCTACAAGGNGTTCATGATGATGTTGACATTATTATTACTGCAA 438

Qy   81 Asp  81
          :::
Db  439 GAA 441
```

RESULT 6
ABQ57053
ID   ABQ57053 standard; cDNA; 717 BP.
XX
AC   ABQ57053;
XX
DT   02-AUG-2002 (first entry)
XX
DE   Human colon cancer related nucleotide sequence SEQ ID NO:748.
XX
KW   Human; colon cancer; cancer; tissue profiling; forensic; mapping;
KW   genetic analysis; diagnostic; antisense therapy; gene; ss.
XX
OS   Homo sapiens.
XX
PN   WO20229086-A2.
XX
PD   11-APR-2002.
XX
PF   02-OCT-2001; 2001WO-US030732.
XX
PR   02-OCT-2000; 2000US-0237271P.
XX
PA   (FARB) BAYER CORP.
XX
PI   Burgess C, Astle JH, Carroll E, Catino TJ, Dwivedi P, Molino GA;
PI   Thiaglingam A, Lewis ME;
XX
DR   WPI; 2002-426115/45.
XX
PT   New isolated nucleic acid that is differentially expressed in cancer
PT   tissues useful for determining the presence of colon cancer in a cell or
PT   tissue type, and in antisense therapy.

Claim 1; Fig 1; 796pp; English.

ABQ5306 to ABQ60787 represent isolated nucleic acids (I) differentially expressed in cancer tissues. ABB78993 to ABB79004 represent proteins encoded by the ABQ60776 to ABQ60787 nucleic acid sequences. (I) can be used in antisense therapy. An antibody immunoreactive with a polypeptide encoded by (I) is useful for detecting cancer in a patient sample, and for detecting the presence or absence of a polynucleotide encoded by a nucleic acid which hybridises to (I) in a cell. A probe/primer derived from (I) can be used for determining the presence of a nucleic acid which hybridises to (I), and for determining the phenotype of cells in a sample of cells from a patient. (I) is useful for determining the presence of colon cancer in a cell or tissue type, for determining the presence or state of other type of cancer, in antisense therapy, to generate macroarrays on a solid surface, to identify a chromosome on which the corresponding gene resides, and in tissue profiling, forensics, genetic analysis, mapping and diagnostic application. (I) can be used to raise antibodies, and to screen for peptide analogues and antagonists

Sequence 717 BP; 192 A; 171 C; 155 G; 180 T; 0 U; 19 Other;

ignment Scores:
ed. No.:   2.34e-52
ore: 404.00
rent Similarity: 100.00%
st Local Similarity: 100.00%
ery Match: 98.54%
:   6

Length: 717
Matches: 80
Conservative: 0
Mismatches: 0
Indels: 0
Gaps: 0

-09-648-310-4 (1-81) x ABQ57053 (1-717)

```
    1 MetAsnValAspHisGluValAsnLeuLeuValGluGlyIleHisArgLeuGlySerLys 20
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
  202 ATGAATGTGGATCACGAGGTTAACCTCTTAGTGGAGGAAATTCATCGTTGGGTTCAAAA 261

   21 AsnAlaAspGlyValLysLeuSerValLysPheGlyValLeuPheArgAspLysCysAla 40
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
  262 AATGGCTGATGGAAAGTTAAGCGTGAAATTTGGGGTCCTCTTCCGTGATGATAAATGTGCC 321

   41 AsnLeuPheGluValGlyThrLeuLysAlaAlaLysArgArgLysIleValThr 60
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||
  322 AACCTCTTTGAAGCATTGGTAGGAACTCTTAAAGCTGCAAAACGAAGGAAGATTGTAACA 381

   61 TyrProGlyGluLeuLeuGlnGlyValHisAspValAspIleIleLeuLeuGln 80
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||
  382 TATCCAGGAGAGCTGCTTCTACAAGGTGTTCATGATGATGTTGACATTATATTACTGCAA 441
```

ABK11085 standard; cDNA; 780 BP.
ABK11085;

05-JUN-2002 (first entry)

Rat Progression Suppressed Gene 13 (rPSGen 13), cDNA.

Rat; Progression Suppressed Gene 13; rPSGen 13; cytostatic; cancer; blood vessel; nasopharyngeal tumour; thyroid tumour; leukaemia; lymphoma; breast; lung; prostate; ovary; colon; gene; ss.

Rattus sp.

Key   Location/Qualifiers
CDS   170..415
             /*tag= a
             /product= "Progression suppressed gene 13 protein"

WO200216419-A2.

28-FEB-2002.

PF 27-AUG-2001; 2001WO-US026795.
XX
PR 25-AUG-2000; 2000US-00643310.
XX
PA (UYCO) UNIV COLUMBIA NEW YORK.
XX
PI Fisher PB, Kang D, Su Z;
XX
DR WPI; 2002-280914/32.
DR P-PSDB; AAU76532.
XX
PT New rat and human Progression Suppressed Gene 13 for preventing the
PT growth of cancer cells and/or new blood vessels, and for treating
PT patients suffering from a cancer.
XX
PS Claim 2; Fig 1; 53pp; English.
XX
CC The invention relates to novel isolated nucleic acids which encode a rat
CC or human Progression Suppressed Gene 13 (PSGen 13) protein. The nucleic
CC acids are useful for preventing the growth of cancer cells and/or new
CC blood vessels and for treating patients suffering from a cancer, e.g.
CC nasopharyngeal tumour, thyroid tumour, leukaemia, lymphoma, or cancer of
CC the breast, lung, prostate, ovary or colon. PSGen 13 may also be used to
CC suppress the transformed phenotype of a malignant cell. Administration of
CC PSGen 13 gene or protein may result in a decrease in tumour mass, number
CC of cancer cells, serum tumour marker, tumour metastasis, vascularisation,
CC perfusion, or rate of tumour growth, improved clinical symptoms, and/or
CC increased patient survival. The present sequence represents the coding
CC sequence of rat Progression Suppressed Gene 13 (rPSGen 13)
XX
SQ Sequence 780 BP; 223 A; 151 C; 187 G; 219 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:   3.83e-50
Score: 390.00
Percent Similarity: 98.77%
Best Local Similarity: 93.83%
Query Match: 95.12%
DB: 6

Length: 780
Matches: 76
Conservative: 4
Mismatches: 1
Indels: 0
Gaps: 0

US-09-648-310-4 (1-81) x ABK11085 (1-780)

```
Qy   1 MetAsnValAspHisGluValAsnLeuLeuValAsnLeuValGluGluIleHisArgLeuGlySerLys 20
       ||||||||:|||||:|||||||||||||||| :||||||||||||||||||||||||
Db 170 ATGAACGTGGACATGAGCTTAACCTCCTGGTGGAGGAAATTCATCGTCCTGGGTTCCAA 229

Qy  21 AsnAlaAspGlyLysLeuSerValLysPheGlyValLeuPheArgAspLysCysAla 40
       |||||||||||:||||||||||||||||||||||||||:||:||||||||||||||||
Db 230 AATGCCGATGGGAAACTGAGTGTGAAGTTTGGGGTCCTCTTCCAAGACGACAGATGTGCC 289

Qy  41 AsnLeuPheGluAlaLeuValGlyThrLeuLysAlaAlaLysArgArgLysIleValThr 60
       |||||||||||:|||||||||||||||||||:|||||||||:||||:||||||||||||
Db 290 AATCTCTTTGAAGCGTTGGTGGGAACTCTGAAAGCCGCAAAACGCGAAGGAGGAGATTGTTACG 349

Qy  61 TyrProGlyGluLeuLeuGlnGlyValHisAspValAspIleIleLeuLeuGln 80
       |||||||||||||||||||||||||||||||||||:||:|||||||:||||||
Db 350 TACGCAGGAGGAGCTGCTTTTGCAAGGTGTTCATGATGATGTTGACATTGTATTGCTGCAA 409

Qy  81 Asp 81
       |||
Db 410 GAT 412
```

RESULT 8
ACH38863
ID ACH38863 standard; cDNA; 458 BP.
XX
AC ACH38863;
XX
DT 13-OCT-2003 (first entry)
XX
DB Human foetal brain cDNA #230.
XX
KW Human; ss; sequencing by hybridisation; SBH; expressed sequence tag; EST;

genome mapping; biodiversity; genetic disorder.

Homo sapiens.

US2003073623-A1.

17-APR-2003.

30-JUL-2001; 2001US-00918995.

30-JUL-2001; 2001US-00918995.

(DRMA/) DRMANAC R T.
(LABA/) LABAT I.
(STAC/) STACHE-CRAIN B.
(DICK/) DICKSON M C.
(JONE/) JONES L W.

Drmanac RT, Labat I, Stache-Crain B, Dickson MC, Jones LW;

WPI; 2003-615964/58.

New polynucleotide sequences obtained from various cDNA libraries, useful
as hybridization probes, as oligomers for PCR, for chromosome and gene
mapping, in the recombinant production of protein, or in generating
antisense DNA or RNA.

Claim 1; SEQ ID NO 26075; 44pp; English.

The invention relates to an isolated polynucleotide comprising any one of
38043 cDNA sequences, appearing as ACH12789-ACH50831, whose sequence was
determined by the technique of SBH (sequencing by hybridisation). Also
included is a purified polypeptide comprising a sequence corresponding to
a reading frame of the novel polynucleotide. The nucleic acid sequences
are useful in diagnostics as expressed sequence tags (EST) for
identifying expressed genes or for physical mapping of the human genome,
in forensics, in assessing biodiversity, or in identifying mutations
responsible for genetic disorders and other traits. The nucleotide
sequences are also useful as hybridisation probes, as oligomers for PCR,
for chromosome and gene mapping, in the recombinant production of
protein, or in generating antisense DNA or RNA. The purified polypeptide
is useful for generating antibodies specific for it. The present sequence
is one of the 38043 isolated cDNA/EST sequences. Note: The sequence data
for this patent did not form part of the printed specification, but was
obtained in electronic format directly from USPTO at
seqdata.uspto.gov/sequence.html?DocID=20030073623

Q Sequence 458 BP; 112 A; 86 C; 149 G; 107 T; 0 U; 4 Other;

Alignment Scores:
Pred. No.:   5.23e-50   Length: 458
Score: 387.00   Matches: 76
Percent Similarity: 100.00%   Conservative: 0
Best Local Similarity: 100.00%   Mismatches: 0
Query Match: 94.39%   Indels: 0
DB: Gaps: 0

US-09-648-310-4 (1-81) x ACH38863 (1-458)

Qy   1 MetAsnValAspHisGluValAsnLeuLeuValGluGluIleHisArgLeuGlySerLys 20
Qy 231 ATGAATGTGGATCACGAGGTTAACCTCTTAGTGGAGGAAATTCATCGTTTGGGTTCAAAA 290

Qy   21 AsnAlaAspGlyLysLeuSerValPheGlyValLeuPheArgAspAspLysCysAla 40
Qy 291 AATGCTGATGGAAAGTTAAGCGTGAAATTTGGGGTCCTCTTCCGTGATGATAAATGTGCC 350

Qy   41 AsnLeuPheGluAlaLeuValGlyThrLeuLysAlaAlaLysArgArgLysIleValThr 60
Qy 351 AACCTCTTTGAAGCATTGGTAGGAACCTTAAAGCTGCAAAACGAAGAAGAGATTGTAACA 410

Qy   61 TyrProGlyGluLeuLeuGlnGlyValHisAspAspValAspIle 76
Qy 421 (partial)

Db   411 TATCCAGGAGAGCTGCTTCTGCAAGGTGTTCATGATGATGTTGACATT 458

RESULT 9
ABT10814/c
ID   ABT10814 standard; cDNA; 653 BP.
XX
AC   ABT10814;
XX
DT   04-DEC-2002 (first entry)
XX
DE   Human breast cancer associated coding sequence SEQ ID NO: 948.
XX
KW   Human; breast specific gene; breast cancer; differential expression;
KW   cytostatic; gene therapy; gene; ss.
XX
OS   Homo sapiens.
XX
PN   WO200259271-A2.
XX
PD   01-AUG-2002.
XX
PF   25-JAN-2002; 2002WO-US002176.
XX
PR   25-JAN-2001; 2001US-0263757P.
PR   25-APR-2001; 2001US-0286090P.
PR   23-MAY-2001; 2001US-0292517P.
XX
PA   (GENE-) GENE LOGIC INC.
XX
PI   Orr MS, Nation M, Diggans JC, Zeng W;
XX
DR   WPI; 2002-674803/72.
XX
PT   Diagnosing breast cancer in a patient comprises detecting the level of
PT   gene expression in cell or tissue samples, where a differential gene
PT   expression is indicative of breast cancer.
XX
PS   Claim 1; SEQ ID NO 948; 260pp + Sequence Listing; English.
XX
CC   The present invention relates to methods of diagnosing breast cancer in a
CC   patient, which comprise detecting the level of expression in a tissue
CC   sample of two or more genes selected from those shown in ABT09867-
CC   ABT11112, where a differential expression of the genes indicates breast
CC   cancer. The methods are useful in diagnosing, treating, detecting the
CC   progression, and in monitoring treatment of breast cancer in patients.
CC   The methods are also useful as a screening tool for agents that modulate
CC   the onset or progression of breast cancer. The breast cancer genes may be
CC   used as diagnostic markers for the prediction or identification of the
CC   malignant state of breast tissue, for confirming the type and progression
CC   of cancer, and for drug screening and assays. The present sequence is a
CC   coding sequence of the invention. Note: The sequence data for this patent
CC   did not form part of the printed specification, but was obtained in
CC   electronic format directly from WIPO at
CC   ftp.wipo.int/pub.published_pct_sequences
XX
SQ   Sequence 653 BP; 213 A; 135 C; 88 G; 217 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:   1.04e-48   Length: 653
Score: 380.00   Matches: 76
Percent Similarity: 95.06%   Conservative: 1
Best Local Similarity: 93.83%   Mismatches: 4
Query Match: 92.68%   Indels: 0
DB: Gaps: 0

US-09-648-310-4 (1-81) x ABT10814 (1-653)

Qy   1 MetAsnValAspHisGluValAsnLeuLeuValGluGluIleHisArgLeuGlySerLys 20
Db 644 ATGAATGTGGATCACGAGGTTAACCTCTTAGTGGAGGAATATTCATCGTTTGGGTTCAAAA 585

Qy   21 AsnAlaAspGlyLysLeuSerValPheGlyValLeuPheArgAspAspLysCysAla 40

```
584 AATGCTGATGGAAGTTAAGCGTGAAAATTGGGGGTCCTCCCGTGATGATAAATGTGCC 525

 41 AsnLeuPheGluAlaLeuValGlyThrLeuLysAlaAlaLysArgArgLysIleValThr 60
    |||||||||||||||||||||||||||||||||||||||||||||||||||||
524 AACCCCTTTGAAGCATTGGTAGGAACTCTTAAAGCTGCAAAACGAAGAAGAGATTGTAACA 465

 61 TyrProGlyGluLeuLeuGlnGlyValGlyValAspAspValAspIleIleLeuLeuGln 80
    |||||||||||||||||||||||||||||||||||||||||
464 TATCCAGGAGAGCTGCTTCTGCAAGGTGTTCATGATGATGTTGACATTATATTACTGCAA 405

 81 Asp 81
    |||
404 GAT 402


SULT 10
221517
AA221517 standard; cDNA; 800 BP.

AA221517;

01-DEC-1999   (first entry)

Progression suppressed gene 13 (PSGen13).

Progression suppressed gene; PSGen; progression elevated gene; PEGen;
tumour; reciprocal subtraction differential RNA display; RSDD;
differential expression; gene cloning; cancer; ss.

Rattus sp.

Key          Location/Qualifiers
CDS          170..448
             /*tag= a
             /product= "PSGen13"
             /note= "Progression suppressed gene 13 protein"

WO994344-A1.

02-SEP-1999.

26-FEB-1999;   99WO-US004323.

27-FEB-1998;   98US-00032684.
03-NOV-1998;   98US-00185115.
23-NOV-1998;   98US-00197889.

(UYCO ) UNIV COLUMBIA NEW YORK.

Fisher PB;

WPI; 1999-550872/46.
P-PSDB; AAY39325.

Identifying nucleic acids differentially expressed between two samples,
particularly sequences involved in tumor progression.

Claim 21; Fig 35B; 110pp; English.

This sequence is the progression suppressed gene 13 (PSGen13). This gene
has suppressed expression in progressed tumour cells. PSGen13 was
identified using new methods for identifying nucleic acids differentially
expressed between two samples. The method involves performing reciprocal
subtraction differential RNA display (RSDD) between the two samples to
generate two subtraction samples. The subtraction samples are amplified
and compared to identify those nucleic acids that are differentially
expressed. The method is used to identify and clone differentially
expressed genes, particularly those with increased or reduced expression
during tumour cell progression, e.g. progression suppressed genes (PSGen)
and progression elevated genes (PEGen). The method reduces the complexity
of the band pattern produced in conventional differential RNA display
(where bands may be obscured, resulting in false positive signals) since
most bands common to both samples are eliminated, allowing identification
and cloning of genes displaying anticipated differential expression. RSDD
```

```
CC   requires only a single anchored primer for amplification and reamplified
CC   cDNA can be analysed by reverse Northern blotting
XX
SQ   Sequence 800 BP; 243 A; 153 C; 185 G; 219 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                        Length:        800
Score:        2.94e-44            Matches:        74
              352.00             Conservative:     4
Percent Similarity:     93.98%   Mismatches:       3
Best Local Similarity:  89.16%   Indels:           2
Query Match:            85.85%   Gaps:             0
DB:           2

US-09-648-310-4 (1-81) x AA221517 (1-800)

Qy   1 MetAsnValAspHisGluValAsnLeuLeuValGlnGluIleHisArgLeuGlySerLys 20
        |||||||||:::|||||||||||||||||||||||||||||||||||||||||
Db 170 ATGAACGTGGACCATGGAGGTTAACCTCCTGGTGGAGCAAATTCATCGTCTGGGTTCCAAA 229

Qy  21 AsnAlaAspGlyLysLeuSerValLysPheGlyValLeuPheArgAspAspLysCysAla 40
        |||||||||||||||||||||||||||||||||||||||||||:::||||||
Db 230 AATGCCGATGGGAAACTGAGTGTGAAGTTTGGGGTCCTCTTCCAAGACGACAGATGTGCC 289

Qy  41 AsnLeuPheGlnAla-LeuValGlyThrLeuLeuLysAla-AlaLysArgArgLysIleValT 60
        ||||||||||||    ||||||||||||||||||||    |||||||||||||||
Db 290 AATCTCTTTGAAACCGTTGGTGGGAACTCTGAAAGCCCGCAAAACGAAGAAGAGATTGTTA 349

Qy  60 hrTyrProGlyGluLeuLeuGlnGlyValHisAspAspValAspIleIleLeuLeuG 80
        |||||||||||||||||||||||||||||||||||:::|||||||||||
Db 350 CGTACGCAGAAGAGCTGCTTTTGCAAGGTGTTCATGATGATGTTGACATTGATTGCTGC 409

Qy  80 lnAsp 81
        |||||
Db 410 AAGAT 414


RESULT 11
AA53910
ID   AA53910 standard; DNA; 551 BP.
XX
AC   AAC53910;
XX
DT   18-OCT-2000 (first entry)
XX
DE   Zea mays DNA fragment SEQ ID NO: 76053.
XX
KW   Hybridisation assay; genetic mapping; gene expression control;
KW   protein identification; signal transduction pathway; metabolic; pathway;
KW   promoter; termination sequence; corn; ss.
XX
OS   Zea mays subsp. mays.
XX
PN   EP1033405-A2.
XX
PD   06-SEP-2000.
XX
PF   25-FEB-2000;   2000EP-00301439.
XX
PR   25-FEB-1999;   99US-0121825P.
PR   05-MAR-1999;   99US-0123180P.
PR   09-MAR-1999;   99US-0123548P.
PR   23-MAR-1999;   99US-0125788P.
PR   25-MAR-1999;   99US-0126264P.
PR   29-MAR-1999;   99US-0126785P.
PR   01-APR-1999;   99US-0127462P.
PR   06-APR-1999;   99US-0128234P.
PR   08-APR-1999;   99US-0128714P.
PR   16-APR-1999;   99US-0129845P.
PR   19-APR-1999;   99US-0130077P.
PR   21-APR-1999;   99US-0130449P.
PR   23-APR-1999;   99US-0130510P.
PR   23-APR-1999;   99US-0130891P.
PR   28-APR-1999;   99US-0131449P.
PR   30-APR-1999;   99US-0132048P.
```

```
30-APR-1999; 99US-0132407P.        PR 21-JUL-1999; 99US-0145088P.
04-MAY-1999; 99US-0132484P.        PR 22-JUL-1999; 99US-0145085P.
05-MAY-1999; 99US-0132485P.        PR 22-JUL-1999; 99US-0145087P.
06-MAY-1999; 99US-0132486P.        PR 22-JUL-1999; 99US-0145089P.
07-MAY-1999; 99US-0132487P.        PR 22-JUL-1999; 99US-0145192P.
11-MAY-1999; 99US-0132663P.        PR 23-JUL-1999; 99US-0145145P.
14-MAY-1999; 99US-0134256P.        PR 23-JUL-1999; 99US-0145218P.
14-MAY-1999; 99US-0134218P.        PR 23-JUL-1999; 99US-0145224P.
14-MAY-1999; 99US-0134219P.        PR 26-JUL-1999; 99US-0145276P.
14-MAY-1999; 99US-0134221P.        PR 27-JUL-1999; 99US-0145913P.
18-MAY-1999; 99US-0134370P.        PR 27-JUL-1999; 99US-0145918P.
18-MAY-1999; 99US-0134768P.        PR 27-JUL-1999; 99US-0145919P.
19-MAY-1999; 99US-0134941P.        PR 28-JUL-1999; 99US-0145951P.
20-MAY-1999; 99US-0135124P.        PR 02-AUG-1999; 99US-0146386P.
21-MAY-1999; 99US-0135353P.        PR 02-AUG-1999; 99US-0146388P.
24-MAY-1999; 99US-0135629P.        PR 03-AUG-1999; 99US-0146389P.
25-MAY-1999; 99US-0136021P.        PR 03-AUG-1999; 99US-0147038P.
27-MAY-1999; 99US-0136392P.        PR 04-AUG-1999; 99US-0147204P.
28-MAY-1999; 99US-0136782P.        PR 04-AUG-1999; 99US-0147302P.
01-JUN-1999; 99US-0137222P.        PR 05-AUG-1999; 99US-0147192P.
03-JUN-1999; 99US-0137523P.        PR 05-AUG-1999; 99US-0147260P.
04-JUN-1999; 99US-0137502P.        PR 06-AUG-1999; 99US-0147303P.
07-JUN-1999; 99US-0137724P.        PR 06-AUG-1999; 99US-0147416P.
08-JUN-1999; 99US-0138094P.        PR 09-AUG-1999; 99US-0147493P.
10-JUN-1999; 99US-0138540P.        PR 09-AUG-1999; 99US-0147935P.
10-JUN-1999; 99US-0138847P.        PR 10-AUG-1999; 99US-0148171P.
14-JUN-1999; 99US-0139919P.        PR 11-AUG-1999; 99US-0148319P.
16-JUN-1999; 99US-0139452P.        PR 12-AUG-1999; 99US-0148341P.
16-JUN-1999; 99US-0139453P.        PR 13-AUG-1999; 99US-0148565P.
17-JUN-1999; 99US-0139492P.        PR 13-AUG-1999; 99US-0148684P.
18-JUN-1999; 99US-0139454P.        PR 16-AUG-1999; 99US-0149368P.
18-JUN-1999; 99US-0139455P.        PR 17-AUG-1999; 99US-0149175P.
18-JUN-1999; 99US-0139456P.        PR 18-AUG-1999; 99US-0149426P.
18-JUN-1999; 99US-0139457P.        PR 18-AUG-1999; 99US-0149722P.
18-JUN-1999; 99US-0139458P.        PR 20-AUG-1999; 99US-0149723P.
18-JUN-1999; 99US-0139459P.        PR 20-AUG-1999; 99US-0149929P.
18-JUN-1999; 99US-0139460P.        PR 23-AUG-1999; 99US-0149902P.
18-JUN-1999; 99US-0139461P.        PR 23-AUG-1999; 99US-0149930P.
18-JUN-1999; 99US-0139462P.        PR 25-AUG-1999; 99US-0150566P.
18-JUN-1999; 99US-0139463P.        PR 26-AUG-1999; 99US-0150084P.
18-JUN-1999; 99US-0139750P.        PR 27-AUG-1999; 99US-0151065P.
18-JUN-1999; 99US-0139763P.        PR 27-AUG-1999; 99US-0151066P.
21-JUN-1999; 99US-0139817P.        PR 30-AUG-1999; 99US-0151080P.
22-JUN-1999; 99US-0139899P.        PR 31-AUG-1999; 99US-0151303P.
23-JUN-1999; 99US-0140353P.        PR 01-SEP-1999; 99US-0151438P.
23-JUN-1999; 99US-0140354P.        PR 07-SEP-1999; 99US-0151930P.
24-JUN-1999; 99US-0140695P.        PR 10-SEP-1999; 99US-0152363P.
28-JUN-1999; 99US-0140823P.        PR 13-SEP-1999; 99US-0153070P.
29-JUN-1999; 99US-0140991P.        PR 15-SEP-1999; 99US-0153758P.
30-JUN-1999; 99US-0141287P.        PR 16-SEP-1999; 99US-0154018P.
01-JUL-1999; 99US-0141842P.        PR 20-SEP-1999; 99US-0154039P.
01-JUL-1999; 99US-0142154P.        PR 22-SEP-1999; 99US-0154779P.
02-JUL-1999; 99US-0142055P.        PR 23-SEP-1999; 99US-0155139P.
06-JUL-1999; 99US-0142390P.        PR 24-SEP-1999; 99US-0155486P.
08-JUL-1999; 99US-0142803P.        PR 28-SEP-1999; 99US-0155559P.
09-JUL-1999; 99US-0142920P.        PR 29-SEP-1999; 99US-0156458P.
12-JUL-1999; 99US-0142977P.        PR 04-OCT-1999; 99US-0156596P.
13-JUL-1999; 99US-0143542P.        PR 05-OCT-1999; 99US-0157117P.
14-JUL-1999; 99US-0143624P.        PR 06-OCT-1999; 99US-0157753P.
15-JUL-1999; 99US-0144005P.        PR 07-OCT-1999; 99US-0157865P.
16-JUL-1999; 99US-0144085P.        PR 08-OCT-1999; 99US-0158029P.
16-JUL-1999; 99US-0144086P.        PR 12-OCT-1999; 99US-0158232P.
19-JUL-1999; 99US-0144325P.        PR 13-OCT-1999; 99US-0158369P.
19-JUL-1999; 99US-0144331P.        PR 13-OCT-1999; 99US-0159293P.
19-JUL-1999; 99US-0144332P.        PR 14-OCT-1999; 99US-0159294P.
19-JUL-1999; 99US-0144333P.        PR 14-OCT-1999; 99US-0159295P.
19-JUL-1999; 99US-0144334P.        PR 14-OCT-1999; 99US-0159329P.
19-JUL-1999; 99US-0144335P.        PR 14-OCT-1999; 99US-0159330P.
20-JUL-1999; 99US-0144352P.        PR 18-OCT-1999; 99US-0159331P.
20-JUL-1999; 99US-0144463P.        PR 21-OCT-1999; 99US-0159637P.
20-JUL-1999; 99US-0144484P.        PR          99US-0159638P.
21-JUL-1999; 99US-0144814P.        PR          99US-0159584P.
21-JUL-1999; 99US-0145086P.        PR          99US-0160741P.
```

XX DR WPI; 2002-304127/34.
XX
PT Identifying a stress condition to which a plant cell has been exposed and
PT producing plants with increased tolerance to these abiotic stresses.
XX
PS Claim 144; SEQ ID NO 355; 577pp + Sequence Listing; English.
XX
CC The invention relates to identifying a stress condition to which a plant
CC cell has been exposed, comprising: (a) contacting nucleic acid
CC representative of expressed polynucleotides in the plant cell with an
CC array or probes representative of the plant cell genome; and (b)
CC detecting a profile of expressed polynucleotides in the plant cell
CC characteristic of a stress response. The method is useful in the
CC production of transgenic plants, cells and seeds and in producing plants
CC with increased tolerance to abiotic stress. The present sequence is that
CC of an Arabidopsis thaliana stress regulated gene (ABZ12196-ABZ17574) used
CC in methods of the invention. Note: The sequence data for this patent is
CC not represented in the printed specification but is based on sequence
CC information supplied to Derwent by the European Patent Office
XX
SQ Sequence 486 BP; 142 A; 103 C; 116 G; 125 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:   2.26e-34   Length: 486
Score: 286.00   Matches: 55
Percent Similarity: 80.77%   Conservative: 8
Best Local Similarity: 70.51%   Mismatches: 15
Query Match: 69.76%   Indels: 0
DB: Gaps: 0

US-09-648-310-4 (1-81) x ABZ12550 (1-486)

Qy   1 MetAsnValAspHisGluValAsnLeuValGluGluIleHisArgLeuGlySerLys 20
        ||||||||||||||||||||||||||||||::::|||||||||||||||||||
Db 199 ATGAACGTAGACGAAGAGATTCAGAAACTGAAGAAGAGATCCATCGTCTTGGTTCTTCGT 258

Qy   21 AsnAlaAspGlyLysLeuSerValLysPheGlyValLeuPheArgAspAspLysCysAla 40
        ||||||||||||||||||||||||||::::::||||||||||||||
Db 259 CAGACCGATGGTCTTCAAGGTGACGTTTGGAGTGTTGTTCAATGACGATCGATGTGCC 318

Qy   41 AsnLeuPheGluAlaLeuValGlyThrLeuLysAlaAlaLysArgLysIleValThr 60
        |||::::||||||||||||||||::::::|||||||||
Db 319 AACACTTTGAAGCATTGGTTGGGACTCTGAGACCTGCCAAGAAACCAAAATAGTCCCA 378

Qy   61 TyrProGlyGluLeuLeuGlnGlyValHisAspValAspIleIleLeu 78
        |||::||||||||||||||||||||
Db 379 TTTGAAGGTGAACTTTTGCTTCAAGGCGTTCACGATAAGGTTGAGATCACTCTC 432

RESULT 13
ADB1847/c
ID ADB1847 standard; cDNA; 492 BP.
XX
AC ADB1847;
XX
DT 29-JAN-2004 (first entry)
XX
DE Arabidopsis thaliana expressed polynucleotide seq id 618.
XX
KW genetically modified organism; transgenic organism; plant;
KW inhibitor testing; activator testing; modifier testing; fungicide;
KW insecticide; genetic function; genetic regulation; cellular metabolism;
KW gene; ss.
XX
OS Arabidopsis thaliana.
XX
PN US2003115639-A1.
XX
PD 19-JUN-2003.
XX
PF 26-JAN-2001; 2001US-00770961.
XX
PR 27-JAN-2000; 2000US-0178466P.

```
(GORL/)  GORLACH J.
(ANYY/)  AN Y.
(HAMI/)  HAMILTON C M.
(PRIC/)  PRICE J L.
(RAIN/)  RAINES T M.
(YUYY/)  YU Y.
(RAME/)  RAMEAKA J G.
(PAGE/)  PAGE A.
(MATH/)  MATHEW A V.
(LEDF/)  LEDFORD B L.
(WOES/)  WOESSNER J P.
(HAAS/)  HAAS W D.
(GARC/)  GARCIA C A.
(KRIC/)  KRICKER M.
(SLAT/)  SLATER T.
(DAVI/)  DAVIS K R.
(ALLE/)  ALLEN K.
(HOFF/)  HOFFMAN N.
(HURB/)  HURBAN P.

Gorlach J,  An Y,  Hamilton CM,  Price JL,  Raines TM,  Yu Y;
Rameaka JG,  Page A,  Mathew AV,  Ledford BL,  Woessner JP,  Haas WD;
Garcia CA,  Kricker M,  Slater T,  Davis KR,  Allen K,  Hoffman N;
Hurban P;

WPI; 2003-810930/76.

Novel Arabidopsis thaliana nucleic acids useful for generating
genetically modified transgenic organisms, for screening biologically
active agents such as fungicides, insecticides.

Claim 1; SEQ ID NO 618; 44pp; English.

The invention describes a nucleic acid (I) comprising a sequence capable
of hybridising under stringent conditions to any one of 999 fully defined
Arabidopsis thaliana sequences (1) as given in specification e.g. 360,
1137, 455, 219, 472 nucleotides, etc. or its fragment. (I) is useful as a
hybridisation probe to complementary molecules in a cDNA library. (I) is
also useful for generating genetically modified and transgenic organisms,
usually plant cells and plants. A protein encoded by (I) is useful in
screening assays to determine the effect of candidate inhibitors,
activators or modifiers of the gene product. The protein is also useful
for screening biologically active agents e.g., fungicides and
insecticides. A genetically modified cell, comprising an exogenous
nucleic acid, where the nucleic acid comprises transcription regulatory
sequences operably linked to a sequence capable of hybridising under
stringent conditions to (1) is useful in the study of genetic function
and regulation, for alteration of the cellular metabolism and for
screening compounds that may affect the biological function of the gene
or gene product. This sequence represents an Arabidopsis thaliana
polynucleotide of the invention.
```

X

Q Sequence 492 BP; 137 A; 103 C; 101 G; 151 T; 0 U; 0 Other;

```
lignment Scores:
red. No.:              2.3e-34       Length:          492
core:                  286.00        Matches:          55
ercent Similarity:     80.77%        Conservative:      8
est Local Similarity:  70.51%        Mismatches:       15
uery Match:            69.76%        Indels:            0
B:                     9             Gaps:              0
```

S-05-648-310-4 (1-81) x ADE81847 (1-492)

```
Y      1 MetAsnValAspHisGluValAsnLeuLeuValGluGluIleHisArgLeuGlySerLys 20
           |||       |||        |||    |||    |||:::    ||::::
b    457 ATGAACGTAGACGAAGAGATTCAGAAACTGAAGAAGAGATCCATCGTCTTGGTTCTCGT 398

Y     21 AsnAlaAspGlyLysLeuSerValLysPheGluGlyValLeuPheArgAspAspLysCysAla 40
           |||    |||       |||         |||:::||:::|||||  |::||   |||
b    397 CAGACCGATGGCGTCTTACAAGGTGACGTTTGGAGTTGTTGTTCAATGACGATCGATGTGCC 338
```

```
Qy     41 AsnLeuPheGluAlaLeuValGlyThrLeuLysAlaAlaLysArgArgLysIleValThr 60
             |||:::|||::||||   ::||:::::|||||||||||||:::||||||:::
Db    337 AACATCTTTGAAGCATTGGTTGGGACTCTGAAGCTGCCAAGAAACGCAAAATAGTCGCA 278

Qy     61 TyrProGlyValLeuLeuLeuGlnGlyValHisAspAspValAspIleIleLeu 78
             |||||||||||||  ||||||   |||||||||||||:::||:::||
Db    277 TTTGAAGGTGAACTTTTGCTTCAGGCGTTCACGATAAGGTTGAGATCCTC 224
```

```
RESULT 14
AAC53673
ID   AAC53673 standard; DNA; 657 BP.
XX
AC   AAC53673;
XX
DT   18-OCT-2000 (first entry)
XX
DE   Arabidopsis thaliana DNA fragment SEQ ID NO: 75315.
XX
KW   Hybridisation assay; genetic mapping; gene expression control;
KW   protein identification; signal transduction pathway; metabolic pathway;
KW   promoter; termination sequence; ss.
XX
OS   Arabidopsis thaliana.
XX
PN   EP1033405-A2.
XX
PD   06-SEP-2000.
XX
PF   25-FEB-2000; 2000EP-00301439.
XX
PR   25-FEB-1999;   99US-0121825P.
PR   05-MAR-1999;   99US-0123180P.
PR   09-MAR-1999;   99US-0123548P.
PR   23-MAR-1999;   99US-0125788P.
PR   25-MAR-1999;   99US-0126264P.
PR   29-MAR-1999;   99US-0126785P.
PR   01-APR-1999;   99US-0127462P.
PR   06-APR-1999;   99US-0127464P.
PR   08-APR-1999;   99US-0128714P.
PR   16-APR-1999;   99US-0129845P.
PR   19-APR-1999;   99US-0130077P.
PR   21-APR-1999;   99US-0130449P.
PR   23-APR-1999;   99US-0130510P.
PR   23-APR-1999;   99US-0130891P.
PR   28-APR-1999;   99US-0131449P.
PR   30-APR-1999;   99US-0132048P.
PR   30-APR-1999;   99US-0132407P.
PR   04-MAY-1999;   99US-0132484P.
PR   05-MAY-1999;   99US-0132485P.
PR   06-MAY-1999;   99US-0132486P.
PR   06-MAY-1999;   99US-0132487P.
PR   07-MAY-1999;   99US-0132863P.
PR   11-MAY-1999;   99US-0134256P.
PR   14-MAY-1999;   99US-0134218P.
PR   14-MAY-1999;   99US-0134219P.
PR   14-MAY-1999;   99US-0134221P.
PR   14-MAY-1999;   99US-0134370P.
PR   18-MAY-1999;   99US-0134768P.
PR   19-MAY-1999;   99US-0134941P.
PR   20-MAY-1999;   99US-0135124P.
PR   21-MAY-1999;   99US-0135353P.
PR   24-MAY-1999;   99US-0135629P.
PR   25-MAY-1999;   99US-0136021P.
PR   27-MAY-1999;   99US-0136392P.
PR   28-MAY-1999;   99US-0136782P.
PR   01-JUN-1999;   99US-0137222P.
PR   03-JUN-1999;   99US-0137528P.
PR   04-JUN-1999;   99US-0137502P.
PR   07-JUN-1999;   99US-0137724P.
PR   08-JUN-1999;   99US-0138094P.
PR   10-JUN-1999;   99US-0138540P.
PR   10-JUN-1999;   99US-0138847P.
PR   14-JUN-1999;   99US-0139119P.
```

Reference	Date		PR Reference	Date
99US-0139452P.	16-JUN-1999;		PR 99US-0148341P.	12-AUG-1999;
99US-0139453P.	16-JUN-1999;		PR 99US-0148565P.	13-AUG-1999;
99US-0139492P.	17-JUN-1999;		PR 99US-0148684P.	13-AUG-1999;
99US-0139454P.	18-JUN-1999;		PR 99US-0149368P.	16-AUG-1999;
99US-0139455P.	18-JUN-1999;		PR 99US-0149175P.	17-AUG-1999;
99US-0139456P.	18-JUN-1999;		PR 99US-0149426P.	18-AUG-1999;
99US-0139457P.	18-JUN-1999;		PR 99US-0149722P.	20-AUG-1999;
99US-0139458P.	18-JUN-1999;		PR 99US-0149723P.	20-AUG-1999;
99US-0139459P.	18-JUN-1999;		PR 99US-0149929P.	20-AUG-1999;
99US-0139460P.	18-JUN-1999;		PR 99US-0149902P.	23-AUG-1999;
99US-0139461P.	18-JUN-1999;		PR 99US-0149930P.	23-AUG-1999;
99US-0139462P.	18-JUN-1999;		PR 99US-0150566P.	25-AUG-1999;
99US-0139463P.	18-JUN-1999;		PR 99US-0150884P.	26-AUG-1999;
99US-0139750P.	18-JUN-1999;		PR 99US-0151065P.	27-AUG-1999;
99US-0139763P.	18-JUN-1999;		PR 99US-0151066P.	27-AUG-1999;
99US-0138817P.	21-JUN-1999;		PR 99US-0151080P.	30-AUG-1999;
99US-0139899P.	22-JUN-1999;		PR 99US-0151303P.	31-AUG-1999;
99US-0140353P.	23-JUN-1999;		PR 99US-0151438P.	01-SEP-1999;
99US-0140354P.	23-JUN-1999;		PR 99US-0151930P.	07-SEP-1999;
99US-0140695P.	24-JUN-1999;		PR 99US-0152363P.	10-SEP-1999;
99US-0140823P.	28-JUN-1999;		PR 99US-0153070P.	13-SEP-1999;
99US-0140991P.	29-JUN-1999;		PR 99US-0153758P.	15-SEP-1999;
99US-0141287P.	30-JUN-1999;		PR 99US-0154018P.	16-SEP-1999;
99US-0141842P.	01-JUL-1999;		PR 99US-0154039P.	20-SEP-1999;
99US-0142154P.	01-JUL-1999;		PR 99US-0154779P.	22-SEP-1999;
99US-0142055P.	02-JUL-1999;		PR 99US-0155139P.	23-SEP-1999;
99US-0142390P.	06-JUL-1999;		PR 99US-0155486P.	24-SEP-1999;
99US-0142803P.	08-JUL-1999;		PR 99US-0155659P.	28-SEP-1999;
99US-0142920P.	09-JUL-1999;		PR 99US-0156458P.	29-SEP-1999;
99US-0142977P.	12-JUL-1999;		PR 99US-0156596P.	04-OCT-1999;
99US-0143542P.	13-JUL-1999;		PR 99US-0157117P.	05-OCT-1999;
99US-0143624P.	14-JUL-1999;		PR 99US-0157753P.	06-OCT-1999;
99US-0144005P.	15-JUL-1999;		PR 99US-0157865P.	07-OCT-1999;
99US-0144085P.	16-JUL-1999;		PR 99US-0158029P.	08-OCT-1999;
99US-0144086P.	16-JUL-1999;		PR 99US-0158232P.	12-OCT-1999;
99US-0144325P.	19-JUL-1999;		PR 99US-0158369P.	13-OCT-1999;
99US-0144331P.	19-JUL-1999;		PR 99US-0159293P.	13-OCT-1999;
99US-0144332P.	19-JUL-1999;		PR 99US-0159294P.	13-OCT-1999;
99US-0144333P.	19-JUL-1999;		PR 99US-0159295P.	14-OCT-1999;
99US-0144334P.	19-JUL-1999;		PR 99US-0159329P.	14-OCT-1999;
99US-0144335P.	19-JUL-1999;		PR 99US-0159330P.	14-OCT-1999;
99US-0144352P.	20-JUL-1999;		PR 99US-0159331P.	14-OCT-1999;
99US-0144632P.	20-JUL-1999;		PR 99US-0159637P.	14-OCT-1999;
99US-0144884P.	20-JUL-1999;		PR 99US-0159638P.	14-OCT-1999;
99US-0144814P.	21-JUL-1999;		PR 99US-0159584P.	18-OCT-1999;
99US-0145086P.	21-JUL-1999;		PR 99US-0160741P.	21-OCT-1999;
99US-0145088P.	21-JUL-1999;		PR 99US-0160767P.	21-OCT-1999;
99US-0145085P.	22-JUL-1999;		PR 99US-0160768P.	21-OCT-1999;
99US-0145087P.	22-JUL-1999;		PR 99US-0160770P.	21-OCT-1999;
99US-0145089P.	22-JUL-1999;		PR 99US-0160814P.	21-OCT-1999;
99US-0145192P.	22-JUL-1999;		PR 99US-0160815P.	21-OCT-1999;
99US-0145145P.	23-JUL-1999;		PR 99US-0160980P.	22-OCT-1999;
99US-0145218P.	23-JUL-1999;		PR 99US-0160981P.	22-OCT-1999;
99US-0145224P.	23-JUL-1999;		PR 99US-0160989P.	22-OCT-1999;
99US-0145276P.	26-JUL-1999;		PR 99US-0161404P.	25-OCT-1999;
99US-0145913P.	27-JUL-1999;		PR 99US-0161405P.	25-OCT-1999;
99US-0145918P.	27-JUL-1999;		PR 99US-0161406P.	25-OCT-1999;
99US-0145919P.	27-JUL-1999;		PR 99US-0161359P.	26-OCT-1999;
99US-0145951P.	28-JUL-1999;		PR 99US-0161360P.	26-OCT-1999;
99US-0146386P.	02-AUG-1999;		PR 99US-0161361P.	26-OCT-1999;
99US-0146388P.	02-AUG-1999;		PR 99US-0161920P.	28-OCT-1999;
99US-0146389P.	02-AUG-1999;		PR 99US-0161992P.	28-OCT-1999;
99US-0147204P.	04-AUG-1999;		PR 99US-0161993P.	28-OCT-1999;
99US-0147302P.	04-AUG-1999;		PR 99US-0162142P.	29-OCT-1999;
99US-0147192P.	05-AUG-1999;			
99US-0147260P.	05-AUG-1999;			
99US-0147303P.	06-AUG-1999;			
99US-0147416P.	06-AUG-1999;			
99US-0147493P.	09-AUG-1999;			
99US-0147595P.	09-AUG-1999;			
99US-0147703P.	10-AUG-1999;			
99US-0148317P.	11-AUG-1999;			
99US-0148319P.	11-AUG-1999;			

Alignment Scores:

Pred. No.:	3.47e-34	Length:	657
Score:	286.00	Matches:	55
Percent Similarity:	80.77%	Conservative:	8
Best Local Similarity:	70.51%	Mismatches:	15
Query Match:	69.76%	Indels:	0
DB:	3	Gaps:	0

```
Score:                 252.00          Matches:       50
Percent Similarity:    100.00%         Conservative:   0
Best Local Similarity: 100.00%         Mismatches:     0
Query Match:            61.46%         Indels:         0
DB:                                     Gaps:           0

US-09-648-310-4 (1-81) x ABK45388 (1-552)

Qy    32 ValLeuPheArgAspAspLysCysAlaAsnLeuPheGluAlaLeuValGlyThrLeuLys 51
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   552 GTCCTCTTCCGTGATGATAAATGTGCCAACCTCTTTGAAGCATTGGTAGGAACTCTTAAA 493

Qy    52 AlaAlaLysArgArgLysIleValThrTyrProGlyGluLeuLeuGlnGlyValHis 71
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   492 GCTGCAAAACGAAGGAAGATTGTAACATATCCAGGAGAGCTGCTTCTGCAAGGTGTTCAT 433

Qy    72 AspAspValAspIleIleLeuLeuGlnAsp 81
            ||||||||||||||||||||||||||||||
Db   432 GATGATGTTGACATTATATTACTGCAAGAT 403

Search completed: March 17, 2004, 14:37:13
Job time : 322 secs
```

```
5-09-648-310-4 (1-81) x AAC53673 (1-657)

     1 MetAsnValAspHisGluValAsnLeuLeuValGluGluIleHisArgLeuGlySerLys 20
         ||||||||||||       |||::::            :: ||      :||
   199 ATGAACGTAGACGAAGAAGATTCAGAAACTGGAAGAAGATCCATCGTCCTTGGTTCCGT 258

    21 AsnAlaAspGlyLysLeuSerValLysPheGlyValLeuPheArgAspAspLysCysAla 40
         ||||                |||          |||:::||||||||
   259 CAGACCGATGGCGTCTTACAAGGTGACGTTGGAGTGTTGTTCAATGACGATCGATGTGCC 318

    41 AsnLeuPheGluAlaLeuValGlyThrLeuLysAlaAlaLysArgArgLysIleValThr 60
         ||||||||||||||||||||||||    |::||||||||||||||    ||
   319 AACATCTTTGAAGCATTGGTAGGGACTCTGAAGACCTGCCAAGAAACGCAAAATAGTCCA 378

    61 TyrProGlyGluLeuLeuGlnGlyValHisAspAspValAspIleIleLeu 78
         ::: ||||||||||||||||||||||||||||||||||:::||||||
   379 TTTGAAGGTGAACTTTGCTTCAGGGCGTTCACGATAAGGTTGAGATCACTCTC 432

RESULT 15
BK45388/c
ABK45388 standard; cDNA; 552 BP.

ABK45388;

05-JUN-2002   (first entry)

cDNA encoding colon tumour protein, SEQ ID No 939.

Human; colon tumour; vaccine; colon cancer; immunogenic; immunotherapy;
gene; ss.

Homo sapiens.

WO20021328-A2.

14-FEB-2002.

31-JUL-2001; 2001WO-US024218.

03-AUG-2000; 2000US-0223283P.
28-MAR-2001; 2001US-0279763P.
29-JUN-2001; 2001US-0302051P.

(CORI-) CORIXA CORP.

King GE, Meagher MJ, Xu J,   Secrist H;

WPI; 2002-241739/29.

New colon cancer polypeptides and polynucleotides, useful as vaccines,
for diagnosing, preventing, and treating colon cancer, and as markers for
the progression of cancer.

Claim 1; SEQ ID NO 939; 147pp; English.

The invention relates to polynucleotides encoding colon tumour proteins.
The polynucleotides and encoded polypeptides are useful in pharmaceutical
compositions, such as vaccines, for the diagnosis, prevention, and
treatment of colon cancer. Polynucleotide sequences may be used as
hybridisation probes or primers, and in the design and preparation of
ribozyme molecules for inhibiting expression of tumour polypeptides and
proteins in tumour cells. The compositions are useful for stimulating an
immune response against cancer, particularly for the immunotherapy of
colon cancer, and as markers for the progression of cancer. ABK44450-
ABK46237 represent coding sequences of human colon tumour proteins of the
invention. Note: With the exception of SEQ ID No 1 and 2, the sequence
data for this patent did not form part of the printed specification but
was supplied by the European Patent Office

Sequence 552 BP; 190 A; 104 C; 75 G; 183 T; 0 U; 0 Other;

lignment Scores:
red. No.:          4.84e-29     Length:      552
```

```
          GenCore version 5.1.6
     Copyright (c) 1993 - 2004 Compugen Ltd.

nucleic - nucleic search, using sw model

n on:     March 17, 2004, 11:45:37 ;  Search time 3704 Seconds
                                       (without alignments)
                                       9770.905 Million cell updates/sec

:le:      US-09-648-310-3
:fect score: 835
quence:   1 ggcacgaggcttgagcgcag.........taaagacactgtaactgtctt 835

oring table:  IDENTITY_NUC
              Gapop 10.0 , Gapext 1.0

arched:   3470272 segs, 21671516995 residues

:al number of hits satisfying chosen parameters:      6940544

nimum DB seq length: 0
nimum DB seq length: 2000000000
st-processing: Minimum Match 0%
               Maximum Match 100%
               Listing first 45 summaries

:abase :  GenEmbl:*
          1:  gb_ba:*
          2:  gb_htg:*
          3:  gb_in:*
          4:  gb_om:*
          5:  gb_ov:*
          6:  gb_pat:*
          7:  gb_ph:*
          8:  gb_pl:*
          9:  gb_pr:*
         10:  gb_ro:*
         11:  gb_sts:*
         12:  gb_sy:*
         13:  gb_un:*
         14:  gb_vi:*
         15:  em_ba:*
         16:  em_fun:*
         17:  em_hum:*
         18:  em_in:*
         19:  em_mu:*
         20:  em_om:*
         21:  em_ov:*
         22:  em_pat:*
         23:  em_ph:*
         24:  em_pl:*
         25:  em_ro:*
         26:  em_sts:*
         27:  em_un:*
         28:  em_vi:*
         29:  em_htg_hum:*
         30:  em_htg_inv:*
         31:  em_htg_other:*
         32:  em_htg_mus:*
         33:  em_htg_pln:*
         34:  em_htg_rod:*
         35:  em_htg_mam:*
         36:  em_htg_vrt:*
         37:  em_sy:*
         38:  em_htgo_hum:*
         39:  em_htgo_mu:*
         40:  em_hcgo_other:*

Pred. No. is the number of results predicted by chance to have a
```

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Match Length	DB	ID	Description
1		835	100.0	835	6	AX456992	AX456992 Sequence
2		835	100.0	835	6	AX456995	AX456995 Sequence
3		762.6	91.3	876	17	AF116682	Af116682 Homo sapi
4		747	89.5	786	6	AR379534	AR379534 Sequence
5		698	83.6	743	9	BC014953	BC014953 Homo sapi
6		578.4	69.3	100296	9	AL590308	AL590308 Human DNA
7		542.8	65.0	552	6	AX396724	AX396724 Sequence
8	c	382.4	45.8	406	6	AX408040	AX408040 Sequence
9	c	365.2	43.7	368	6	AX341854	AX341854 Sequence
10	c	335.2	40.1	780	6	AX456990	AX456990 Sequence
11		335.2	40.1	780	6	AX456994	AX456994 Sequence
12	c	312.8	37.5	692	10	AF065991	AF065991 Mus muscu
13		279.6	33.5	296	6	AX918409	AX918409 Sequence
14		279.6	33.5	296	6	BD053942	BD053942 Sequence
15		249.2	29.8	526	10	BC039801	BC039801 Mus muscu
16		244.2	29.2	176094	2	AC102256	AC102256 Mus muscu
17	c	244.2	29.2	198132	2	AC102536	AC102536 Mus muscu
18	c	237.6	28.5	170882	2	AC107839	AC107839 Mus muscu
19		237.6	28.5	219661	10	AC117670	AC117670 Mus muscu
20		226.2	27.1	198793	2	AC117858	AC117858 Rattus no
21		226.2	27.1	239053	2	AC129055	AC129055 Rattus no
22	c	215.2	25.8	231507	2	AC137434	AC137434 Rattus no
23		186	22.3	219	6	AR426459	AR426459 Sequence
24		186	22.3	219	6	BD122012	BD122012 EST and e
25		177	21.2	892	5	AY353856	AY353856 Coturnix
26		170.2	20.4	611	6	BD229701	BD229701 Human gen
27		135.2	16.2	1315	5	BC053269	BC053269 Danio rer
28		133.2	16.0	1283	5	AY398399	AY398399 Danio rer
29		126.8	15.2	114041	2	AP001868	AP001868 Homo sapi
30	c	126.8	15.2	165913	9	AC079616	AC079616 Homo sapi
31	c	126.8	15.2	242184	2	AC015899	AC015899 Homo sapi
32		109.4	13.1	486	6	AX505660	AX505660 Sequence
33		109.4	13.1	516	8	AY087101	AY087101 Arabidops
34	c	97.4	11.7	239874	5	BX248390	BX248390 Zebrafish
35		96.2	11.5	543	8	AK062756	AK062756 Oryza sat
36		91.4	10.9	464	8	SL293059	Z93059 S.latifolia
37		82.6	9.9	98124	8	AT16L1	AT16L1 Arabidops
38	c	82.6	9.9	199536	8	ATCHRIV79	ATCHRIV79 Arabidops
39	c	78.4	9.4	79575	8	AP006357	AP006357 Lotus cor
40	c	71	8.5	636	11	G96685	G96685 S210P4FF1l.
41	c	70.6	8.5	150447	9	AC097277	AC097277 Oryza sat
42		69	8.3	190	6	E25942	E25942 Blastocyst
43		69	8.3	190	6	E25963	E25963 Blastocyst
44	c	55.6	6.7	257757	3	AE014837	AE014837 Plasmodiu
45		55	6.6	349980	6	AX344551	AX344551 Sequence

ALIGNMENTS

```
RESULT 1
AX456992
LOCUS        AX456992                  835 bp        DNA        linear        PAT  06-JUL-2002
DEFINITION   Sequence 3 from Patent WO0216419.
ACCESSION    AX456992
VERSION      AX456992.1  GI:21715784
KEYWORDS     .
SOURCE       Homo sapiens (human)
  ORGANISM   Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE    1
  AUTHORS    Fisher,P.B., Kang,D.C. and Su,Z.Z.
  TITLE      Progression suppressed gene 13 (psgen 13) and uses thereof
  JOURNAL    Patent: WO 0216419-A 3 28-FEB-2002;
```

ACCESSION   AX456995
VERSION   AX456995.1 GI:21715786
KEYWORDS   .
SOURCE   Homo sapiens (human)
ORGANISM   Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
AUTHORS   Fisher,P.B., Kang,D.C. and Su,Z.Z.
TITLE   Progression suppressed gene 13 (psgen 13) and uses thereof
JOURNAL   Patent: WO 0216419-A 6 28-FEB-2002;
             The Trustees of Columbia University in the City of New York (US)
FEATURES   Location/Qualifiers
     source   1..835
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

Query Match   100.0%; Score 835; DB 6; Length 835;
Best Local Similarity   100.0%; Pred. No. 1.5e-179;
Matches 835; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy   1 GGCACGAGGCTTGAGCGCAGAAACACTTACTTTTCCCCTACCCTGCTCCTCCTCCTCCA 60
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 GGCACGAGGCTTGAGCGCAGAAACACTTACTTTTCCCCTACCCTGCTCCTCCTCCTCCA 60

Qy   61 CAGCCGTCTTTCTTTGCCTCAGCCACTTCCTTCCTTCCTTCGCCTCACCCTCACCCAGTGCAC 120
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 CAGCCGTCTTTCTTTGCCTCAGCCACTTCCTTCCTTCCTTCGCCTCACCCTCACCCAGTGCAC 120

Qy 121 TGAAGAAGTAACCGGGTCCAGACCCACCGCGCGCCAGTTCTCCGGCGGGAAGGAAAACC 180
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 121 TGAAGAAGTAACCGGGTCCAGACCCACCGCGCGCCAGTTCTCCGGCGGGAAGGAAAACC 180

Qy 181 GCGCAGAGAGGCAGCAATGAATGTGGATCACGAGGTTAACCTCTTAGTGGAGGAAATTCA 240
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 181 GCGCAGAGAGGCAGCAATGAATGTGGATCACGAGGTTAACCTCTTAGTGGAGGAAATTCA 240

Qy 241 TCGTTTGGGTTCAAAAAATGCTGATGAAAGTTAAGCGTGAAATTTGGGGTCCTCTTCCG 300
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 241 TCGTTTGGGTTCAAAAAATGCTGATGAAAGTTAAGCGTGAAATTTGGGGTCCTCTTCCG 300

Qy 301 TGATGATAAATGTGCCAACCTCTTGAAGCATTGGTAGGAACTCTTAAAGCTGCAAAACG 360
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 301 TGATGATAAATGTGCCAACCTCTTGAAGCATTGGTAGGAACTCTTAAAGCTGCAAAACG 360

Qy 361 AAGGAAGATTGTAACATATCCAGGAGAAGTGTTCTGCAAGGTGTTCATGATGATGTTGA 420
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 361 AAGGAAGATTGTAACATATCCAGGAGAAGTGTTCTGCAAGGTGTTCATGATGATGTTGA 420

Qy 421 CATTATATTACTGCAAGATTAATGTGGTTTACATATCTTTATGTACTGCCATTTTTTGTT 480
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 421 CATTATATTACTGCAAGATTAATGTGGTTTACATATCTTTATGTACTGCCATTTTTTGTT 480

Qy 481 TCTGGTAAACTGGAATATAAAGTGAAAGAACAAACATTGAACATACTTAATGTATTTTT 540
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 481 TCTGGTAAACTGGAATATAAAGTGAAAGAACAAACATTGAACATACTTAATGTATTTTT 540

Qy 541 ATAGAACTTTGTAAACGAAAGGAGATTCATGTTTAGAAGTCTGCCTTTTTATATCTT 600
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 541 ATAGAACTTTGTAAACGAAAGGAGATTCATGTTTAGAAGTCTGCCTTTTTATATCTT 600

Qy 601 GAAAGAAAATCATCATGTATGATGCTATAAAATAAATCCTATTATTTTCTCAGGAATCTGG 660
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 601 GAAAGAAAATCATCATGTATGATGCTATAAAATAAATCCTATTATTTTCTCAGGAATCTGG 660

Qy 661 TTAGGAATTGCAGGCAATGACAGAGATTTTTGCGGGGCAGGGATGGGAATGTTTGTTCATAAA 720
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 661 TTAGGAATTGCAGGCAATGACAGAGATTTTTGCGGGGCAGGGATGGGAATGTTTGTTCATAAA 720

Qy 721 TAATTAGACATTTTCTATAGATATTTGACATTCTGCGAAAGCAACAAGCAAACTGAAGAC 780

```
127  AGGTAACCGGGTCCAGACCGGGGCCAGTTCTCCGGGAGGAAGCGGAGGAAAACCGCGCAG  186
     |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
66   AGGTAACCGGGTCCAGACCGGGGCCAGTTCTCCGGGAGGAAGCGGAGGAAAACCGCGCAG  125

187  AGAGGCAGCAGCAATGAATGTGGATCACGAGGTTAACCTCTTAGTTGGAGGAAATTCATCGTTT  246
     |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
126  AGAGGCAGCAGCAATGAATGTGGATCACGAGGTTAACCTCTTAGTTGGAGGAAATTCATCGTTT  185

247  GGGTTCAAAAAATGCTGATGGAAAGTTAAGCTGAAATTTGGGGTCCTCTTCCGTGATGA  306
     ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
186  GGGTTCAAAAAATGCTGATGGAAAGTTAAGCTGAAATTTGGGGTCCTCTTCCGTGATGA  245

307  TAAATGTGCCAACCTCTTGAAGCATTGGTAGGAACTCTTAAAGCTGCAAAACGAAGGAA  366
     ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
246  TAAATGTGCCAACCTCTTGAAGCATTGGTAGGAACTCTTAAAGCTGCAAAACGAAGGAA  305

367  GATTGTAACATATCCAGGAGAGCTGCTTCGCAAGGTGTTCATGATGGTTGACATTAT  426
     ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
306  GATTGTAACATATCCAGGAGAGCTGCTTCGCAAGGTGTTCATGATGATGGTTGACATTAT  365

427  ATTACTGCAAGATTAATGTGGTTTACATATCTTTATGTACTGCGCCATTTTTGTTTCTGGT  486
     ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
366  ATTACTGCAAGATTAATGTGGTTTACATATCTTTATGTACTGCGCCATTTTTGTTTCTGGT  425

487  AAACTGGAATATAAAGTGAAAGAACAAAACATTGAACATCTTAATGTGATTTTTATAGAA  546
     ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
426  AAACTGGAATATAAAGTGAAAGAACAAAACATTGAACATCTTAATGTGATTTTTATAGAA  485

547  CTTTGTAAACGAAAGGAGATTCATGTTTTAGAAGTCTGTCCTTTTTATATCTTGAAAGA  606
     ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
486  CTTTGTAAACGAAAGGAGATTCATGTTTTAGAAGTCTGTCCTTTTTATATCTTGAAAGA  545

607  AAATCTATGTATGATGCTATAAAATCCTATTATTTTTCTCAGGAATCTCGGTTAGGA  666
     |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
546  AAATCTATGTATGATGCTATAAAATCCTATTATTTTTCTCAGGAATCTCGGTTAGGA  605

667  ATTGCAGGCAATGAGATTTTTGCGGGGATGGGAATGTTGTTCATAATAATTA  726
     |||||||||||||||||||||||||||||||||||||||||||||||||||||
606  ATTGCAGGCAATGAGATTTTTGCGGGGATGGGAATGTTGTTCATAAATAATTA  665

727  GACATTTCTATAGATATTTGACATTCTGCGAAAGCAACAAGCAACTGAAGACCAACTC  786
     |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
666  GACATTTCTATAGATATTTGACATTCTGCGAAAGCAACAAGCAACTGAAGACCAACTC  725

787  CTATGAGAAATATTATGATGTTTATGTAATAAAAGACATGTAACTGTCTT  835
     |||||||||||||||||||||||||||||||||||||||||||||||||
726  CTATGAGAAATATTATGATGTTTATGTAATAAAAGACATGTAACTGTCTT  774
```

```
RESULT 4
AR379534
LOCUS       AR379534             786 bp    DNA        linear    PAT 18-DEC-2003
DEFINITION  Sequence 79 from patent US 6607879.
ACCESSION   AR379534
VERSION     AR379534.1   GI:40087168
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 786)
  AUTHORS   Cocks,B.G., Stuart,S.G. and Seilhamer,J.J.
  TITLE     Compositions for the detection of blood cell and immunological
            response gene expression
  JOURNAL   Patent: US 6607879-A 79 19-AUG-2003;
FEATURES             Location/Qualifiers
     source          1..786
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN

Query Match     89.5%;  Score 747;  DB 6;  Length 786;
Best Local Similarity  99.0%;  Pred. No. 1.6e-159;
Matches  773;  Conservative  0;  Mismatches  5;  Indels  3;  Gaps  2;
```

```
721  TAATTAGACATTTCTATAGAGATATATTGACATTCTGCGAAAGCAACAAGCAACTGAAGAC  780
     ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
781  CAACTCCTATGAGAAATATTATGATGTTTATGTAATAAAAGACATGTAACTGTCTT  835
     |||||||||||||||||||||||||||||||||||||||||||||||||||||||
781  CAACTCCTATGAGAAATATTATGATGTTTATGTAATAAAAGACATGTAACTGTCTT  835
```

```
SULT 3
116682
AF116682   standard; mRNA; HUM; 876 BP.

AF116682;

AF116682.1

24-MAY-2000  (Rel. 63, Created)
09-MAY-2001  (Rel. 67, Last updated, Version 2)

Homo sapiens PRO2013 mRNA, complete cds.

Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.

[1]
1-876
Zhang C., Yu Y., Zhang S., Wei H., Zhou G., Ouyang S., Luo L., Bi J.,
Liu M., He F.;
"Functional prediction of the coding sequences of 121 new genes deduced by
analysis of cDNA clones from human fetal liver";
Unpublished.

[2]
1-876
Zhang C., Yu Y., Zhang S., Wei H., Zhou G., Ouyang S., Luo L., Bi J.,
Liu M., He F.;
;
Submitted (24-DEC-1998) to the EMBL/GenBank/DDBJ databases.
Department of Experimental Hematology, Institute of Radiation Medicine,
Beijing Taiping Road 27, Beijing 100850, P. R. China

SWISS-PROT; Q9P1F3; CFB5_HUMAN.

Key            Location/Qualifiers
source         1..876
               /db_xref="taxon:9606"
               /mol_type="mRNA"
               /organism="Homo sapiens"
               /clone="FLB7547"
               /tissue_type="liver"
               /dev_stage="fetal"
CDS            136..381
               /codon_start=1
               /db_xref="SWISS-PROT:Q9P1F3"
               /note="predicted protein of HQ2013"
               /product="PRO2013"
               /protein_id="AAF71102.1"
               /translation="MNVDHEVNLLVEEIHRLGSKNADGKLSVKFGVLFRDDKCANLFEA
               LVGTLKAAKRRKIVTYPGEELLLQGVHDDVDIILLQD"

Sequence 876 BP; 333 A; 131 C; 167 G; 245 T; 0 other;

Query Match     91.3%;  Score 762.6;  DB 17;  Length 876;
Best Local Similarity  99.5%;  Pred. No. 4.5e-163;
Matches  765;  Conservative  0;  Mismatches  0;  Indels  4;  Gaps  0;

67   TCTTTCTTGCTCAGCCACTTCCTTCCTTCGCCTCACCCTCCCAGTGCACTGCAGGAAGA  126
     |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
6    TTTTTTTTTTGGCTCAGCCACTTCCTTCCTTGGCCTCACCCTCACCCTCCCCAGTGCACTGAAGA  65
```

Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
Fahey,J., Helton,E., Ketteman,M., Madan,A., Rodrigues,S.,
Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E.,
Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.

TITLE   Generation and initial analysis of more than 15,000 full-length
           human and mouse cDNA sequences
JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
MEDLINE   22388257
PUBMED   12477932
REFERENCE 2 (bases 1 to 743)
AUTHORS   Strausberg,R.
TITLE   Direct Submission
JOURNAL   Submitted (01-OCT-2001) National Institutes of Health, Mammalian
           Gene Collection (MGC), Cancer Genomics Office, National Cancer
           Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
           USA
REMARK   NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT   Contact: cgapbs-r@mail.nih.gov
           Email: cgapbs-r@mail.nih.gov
           Tissue Procurement: Dr. Daniel McVicar, DBS/NCI
           cDNA Library Preparation: Rubin Laboratory
           cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
           DNA Sequencing by: Genome Sequence Centre,
           BC Cancer Agency, Vancouver, BC, Canada
           info@bcgsc.bc.ca
           Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield,
           Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin,
           Letticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo
           Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven
           Ness, Pawan Pandoh, Anna-Liisa Prabhu, Parvaneh Saeedi, Jacqueline
           Schein, Duane Smailus, Michael Smith, Lorraine Spence, Jeff Stott,
           Michael Thorne, Miranda Tsai, Natasja van den Bosch, Jill Vardy,
           George Yang, Scott Zuyderduyn, Marco Marra.

           Clone distribution: MGC clone distribution information can be found
           through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
           Series: IRAL Plate: 34 Row: a Column: 8.
FEATURES   Location/Qualifiers
     source   1..743
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="MGC:22978 IMAGE:4849571"
                     /tissue_type="Blood, natural killer cell"
                     /clone_lib="NIH_MGC_106"
                     /lab_host="DH10B-R"
                     /note="Vector: pOTB7"
     gene   1..743
                     /gene="C6orf115"
                     /note="synonyms: PRO2013, HSPC280"
                     /db_xref="LocusID:58527"
     CDS   74..319
                     /codon_start=1
                     /product="C6orf115 protein"
                     /protein_id="AAH14953.1"
                     /db_xref="GI:11528977"
                     /db_xref="LocusID:58527"
                     /translation="MNVDHEVNLLVEEIHRLGSKNADGKLSVKFGVLFRDDKCANLFE
                     ALVGTLKAAKRKIVTYPGELLLQGVHDDVDIILLQD"
ORIGIN

58   CCACAGCCGTCTTCTCTTGCCTCAGCCACTTCCTTCCTTCCTTGCCTCCACCCTCCCAGTG 117
     ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1   CCACAGCCGTCTTTCTCTTGCCTCAGCCACTTCCTTCCTTCCTTGCCTCCACCCTCCCAGTG 60

118 CACTGAAGAAGGTAACCGGGTCCAGACACCCACGGCGCCAGTTCTCCGGCGGAAGGAAA 177
     ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
61   CACTGAAGAAGGTAACCGGGTCCAGACACCCACGGCGCCAGTTCTCCGGCGGGAAGGAAA 120

178 ACCGCGCAGAGAGGCAGCAATGAATGTGATCACGAGGTAACCTCTTAGTGAGGAAAT 237
     ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
121 ACCGCGCAGAGAGGCAGCAATGAATGTGATCACGAGGTTAACCTCTTAGTGAGGAAAT 180

238 TCATCGTTGGGTTCAAAAAATGCTGATGAAAGTAAGC--GTGAAATTTGGGGTCCTC 295
     |||||||||||||||||||||||||||| ||||||| ||||||||||||||||||
181 TCATCGTTTGGGTTCAAAAAATGCTGATGAAAGTTAAGCCGTGAAATTTTGGGGTCCTC 240

296 TTCCGTGATGATAAATGTGCCAACCTCTTTGAAGCATTGGTAGGAACTCTTAAAGCTGC- 354
     ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
241 TTCCGTGATGATAAATGTGCCAACCTCTTTGAAGCATTGGTAGGAACTCTTAAAGCTGCA 300

355 AAAACGAAGAAGAGATTGTAACATATCCAGGAGAGCGTGCTTCTCGCAAGGTGTTCATGATGA 414
     ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
301 AAAACGAAGAAGAGATTGTAACATATCCAGGAGAGCGTGCTTCTCGCAAGGTGTTCATGATGA 360

415 TGTTGACATTATATTACTGCAAGATTAATGTGGTTTACATATCTTTATGTACTGCCATTT 474
     |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
361 TGTTGACATTATATTACTGCAAGATTAATGTGGTTTACATATCTTTATGTACTGCCATTT 420

475 TTTGTTTCTGGTAAACTGGAATATAAAGTGAAAGAACAAACATTGAACATACTTAATGT 534
     ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
421 TTTGTTTCTGGTAAACTGGAATATAAAGTGAAAGAACAAACATTGAACATACTTAATGT 480

535 ATTTTATGAACTTTGTAAACGAAAGGAGATTCATGTTTAGAGAGTCTGTCTCCTTTTTA 594
     ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
481 ATTTTTATGAACTTTGTAAACGAAAGGAGATTCATGTTTAGAGAGTCTGTCTCCTTTTTA 540

595 TATCTTGAAAGAAATCTATGTGATGATGCTATAAATAAATCCTATTATTTTCTCAGGA 654
     ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
541 TATCTTGAAAGAAAATCTATGTGATGATGCTATAAATAAATCCTATTATTTTCTCAGGA 600

655 ATCTGGTTAGGAATTGCAGGCAATGAGATTTTTTGCGGGGCAGGGGATGGGAATGTTGTT 714
     |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
601 ATCTGGTTAGGAATTGCAGGCAATGAGATTTTTTGCGGGGCAGGGGATGGGAATGTTGTT 660

715 CATAAAATAATTAGACATTTCTATAGATATTTGACATTCTCGGAAAGCAACAAGCAAACT 774
     ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
661 CATAAAATAATTAGACATTTTCTATAGATATTTGACATTCTCGGAAAGCAACAAGCAAACT 720

775 GAAGACCAACTCCTATGAGAAAATATTATGATGTTTATGTAATAAAGACATGTAACTGTCT 834
     |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
721 GAAGACCAACTCCTATGAGAAAATATTATGATGTTTATGTAATAAAGACATGTAACTGTCT 780

835 T 835
   |
781 T 781

ESULT 5
C014953
OCUS   BC014953   743 bp   mRNA   linear   PRI 25-NOV-2003
EFINITION Homo sapiens chromosome 6 open reading frame 115, mRNA (cDNA clone
           MGC:22978 IMAGE:4849571), complete cds.
CCESSION   BC014953
ERSION   BC014953.1 GI:11528976
EYWORDS   MGC.
OURCE   Homo sapiens (human)
ORGANISM   Homo sapiens
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
EFERENCE   1 (bases 1 to 743)
AUTHORS   Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
           Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,

```
Query Match          83.6%;  Score 698; DB 9; Length 743;
Best Local Similarity 99.4%;  Pred. No. 2.3e-148;
Matches   712; Conservative   0; Mismatches   0; Indels   4;  Gaps   1;

120 CTGAGAAGGTAACCGGTCCAGACCCACCGGCGCCAGTTCTCCGGCGGAAGGAAAAC 179
    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
  1 CTGAGAAGGTAACCGGTCCAGACCCACCGGCGCCAGTTCTCCGGCGGAAGGAAAAC 60

180 CGGCCAGAGAGGCACGAATGAAATGTGGATCACGAGGTTAACCTCTTAGTGAGGAAATTC 239
    |||||||||||||      |||||||||||||||||||||||||||||||||||||||||
 61 CGGGCAGAG---AGCAATGAATGTGGATCACGAGGTTAACCTCTTAGTGAGGAAATTC 116

240 ATCGTTTGGGTCAAAAAATGCTGATGGAAAGTTAAGCGTGAAATTTGGGGTCCTCTTCC 299
    |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
117 ATCGTTTGGGTTCAAAAAATGCTGATGGAAAGTTAAGCGTGAAATTTGGGGTCCTCTTCC 176

300 GTGATGATAAATGTGCCAACCTCTTTGAAGCATTGGTAGGAACCTCTTAAAGCTGCAAAAC 359
    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
177 GTGATGATAAATGTGCCAACCTCTTTGAAGCATTGGTAGGAACCTCTTAAAGCTGCAAAAC 236

360 GAAGGAAGATTGTAACATATCCAGGAGAGTGCTTCTGCAAGGTGTTCATGATGATGTTG 419
    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
237 GAAGGAAGATTGTAACATATCCAGGAGAGTGCTTCTGCAAGGTGTTCATGATGATGTTG 296

420 ACATTATTATTACTGCAAGATTAAATGTGGTTTACATATCTTTATGTACTGCCATTTTTGT 479
    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
297 ACATTATTATTACTGCAAGATTAATGTGGTTTACATATCTTTATGTACTGCCATTTTTTGT 356

480 TTCTGGTAAACTGGAATATAAAGTGAAAGAACAAACATTTGAACATACTTAATGTATTTT 539
    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
357 TTCTGGTAAACTGGAATATAAAGTGAAAGAACAAACATTTGAACATACTTAATGTATTTT 416

540 TATAGAACTTTGTAAACGAAAGGAGATTCATGTTTTAGAGTCTGCTCCTTTTTATATCT 599
    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
417 TATAGAACTTTGTAAACGAAAGGAGATTCATGTTTTAGAGTCTGCTCCTTTTTATATCT 476

600 TGAAAGAAAATCTATGTATGATGCTATAAAAATAAATCCTATTATTTTTCTCAGGAATCTG 659
    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
477 TGAAAGAAAATCTATGTATGATGCTATAAAAATAAATCCTATTATTTTTCTCAGGAATCTG 536

660 GTTAGGAATTGCAGGCAATGAGTGAGATTTTTTTGCGGGGATGGGAATGTTGTTCATAA 719
    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
537 GTTAGGAATTGCAGGCAATGAGTGAGATTTTTTGCGGGGATGGGAATGTTGTTCATAA 596

720 ATAATTAGACATTTTCTATAGATATTTGACATTCTGCGAAAGCAACAAGCAAACTGAAGA 779
    |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
597 ATAATTAGACATTTTCTATAGATATTTGACATTCTGCGAAAGCAACAAGCAAACTGAAGA 656

780 CCAACTCCTATGAGAAATATTATTGATGTTTATGTAATAAGACATGTAACTGTCTT 835
    |||||||||||||||||||||||||||||||||||||||||||||||||||||||
657 CCAACTCCTATGAGAAATATTATTGATGTTTATGTAATAAAGACATGTAACTGTCTT 712
```

Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e. phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one plasmid subclone or more than one M13 subclone; and the
assembly was confirmed by restriction digest. The following
abbreviations are used to associate primary accession numbers given
in the feature table with their source databases: Em:, EMBL; Sw:,
SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information on the WORMPEP
database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
was generated from part of bacterial clone contigs of human
chromosome 6, constructed by the Sanger Centre Chromosome 6 Mapping
Group. Further information can be found at
http://www.sanger.ac.uk/HGP/Chr6
RP11-501K14 is from the library RPCI-11.2 constructed by the group
of Pieter de Jong. For further details see
http://www.chori.org/bacpac/home.htm
VECTOR: pBACe3.6
IMPORTANT: This sequence is not the entire insert of clone
RP11-501K14 It may be shorter because we sequence overlapping
sections only once, except for a 100 base overlap.
The true left end of clone RP11-501K14 is at 1 in this sequence.
The true left end of clone RP11-225E12 is at 98297 in this sequence.
The true right end of clone RP11-94L3 is at 97304 in this sequence.

FEATURES   Location/Qualifiers
     source   1..100296
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="6"
                     /clone="RP11-501K14"
                     /clone_lib="RPCI-11.2"
     repeat_region   2..99
                     /note="MIR repeat: matches 103. .194 of consensus"
     repeat_region   118..199
                     /note="MER47 repeat: matches 2242. .2323 of consensus"
     repeat_region   190..265
                     /note="MER47 repeat: matches 2242. .2317 of consensus"
     repeat_region   268..346
                     /note="MER47 repeat: matches 1. .79 of consensus"
     repeat_region   502..671
                     /note="MIR repeat: matches 7. .185 of consensus"
     repeat_region   1420..1737
                     /note="AluJb repeat: matches 2. .310 of consensus"
     repeat_region   2010..2033
                     /note="12 copies 2 mer tt 100% conserved"
     repeat_region   2650..2961
                     /note="AluSx repeat: matches 3. .306 of consensus"
     repeat_region   5767..5895
                     /note="MIR repeat: matches 21. .166 of consensus"
     repeat_region   6062..6159
                     /note="L2 repeat: matches 2587. .2709 of consensus"
     repeat_region   6165..6268
                     /note="L2 repeat: matches 2607. .2709 of consensus"
     repeat_region   6916..7157
                     /note="AluJb repeat: matches 1. .295 of consensus"
     repeat_region   7301..7442
                     /note="L2 repeat: matches 1768. .1916 of consensus"
     repeat_region   7746..7999
                     /note="AluSx repeat: matches 45. .298 of consensus"
     repeat_region   8013..8152
                     /note="MIR repeat: matches 14. .151 of consensus"
     repeat_region   8899..8959
                     /note="MIR repeat: matches 194. .256 of consensus"
     repeat_region   9209..9497
                     /note="L1MB8 repeat: matches 5863. .6175 of consensus"

repeat_region   9503..9698
                 /note="AluJo repeat: matches 90..281 of consensus"
repeat_region   9703..10428
                 /note="L1MB8 repeat: matches 5126..5867 of consensus"
repeat_region   10429..10759
                 /note="AluJb repeat: matches 1..310 of consensus"
repeat_region   10760..10773
                 /note="L1MB8 repeat: matches 5114..5126 of consensus"
repeat_region   10897..11050
                 /note="L2 repeat: matches 821..962 of consensus"
repeat_region   11051..11360
                 /note="AluSg repeat: matches 1..310 of consensus"
repeat_region   11361..12691
                 /note="L2 repeat: matches 962..2419 of consensus"
repeat_region   12850..13009
                 /note="AluJb repeat: matches 129..279 of consensus"
repeat_region   15295..15341
                 /note="MIR repeat: matches 97..142 of consensus"
repeat_region   15342..15632
                 /note="AluSg repeat: matches 21..308 of consensus"
repeat_region   15563..15735
                 /note="MIR repeat: matches 142..250 of consensus"
repeat_region   16229..16532
                 /note="AluSx repeat: matches 1..301 of consensus"
repeat_region   16577..16796
                 /note="L2 repeat: matches 1..217 of consensus"
repeat_region   18247..18407
                 /note="MER20 repeat: matches 1..217 of consensus"
repeat_region   18556..18687
                 /note="MER58C repeat: matches 40..88 of consensus"
misc_feature   19512..20572
                 /note="MIR repeat: matches 48..149 of consensus"
misc_feature   20013..20256
                 /note="CpG island"
                 /evidence=not_experimental
repeat_region   20025..20120
                 /note="Sequence from overlapping clone RP11-94L3
                 (AL591033). Assembly confirmed by restriction digest."
repeat_region   20058..20117
                 /note="4 copies 24 mer 72% conserved"
repeat_region   21404..21522
                 /note="20 copies 3 mer gcg 73% conserved"
repeat_region   21843..22026
                 /note="MER46C repeat: matches 82..191 of consensus"
repeat_region   22027..22320
                 /note="L2 repeat: matches 2351..2547 of consensus"
repeat_region   22321..22369
                 /note="AluJb repeat: matches 1..297 of consensus"
repeat_region   22394..22607
                 /note="L2 repeat: matches 2547..2593 of consensus"
repeat_region   22674..22981
                 /note="MIR repeat: matches 32..247 of consensus"
repeat_region   22982..23029
                 /note="AluJb repeat: matches 3..311 of consensus"
repeat_region   23030..23106
                 /note="6 copies 8 mer gaaagaaa 79% conserved"
repeat_region   23146..23213
                 /note="L2 repeat: matches 2667..2749 of consensus"
repeat_region   23216..23317
                 /note="MER6 repeat: matches 799..865 of consensus"
repeat_region   23306..23808
                 /note="LTR38 repeat: matches 3..98 of consensus"
repeat_region   23810..24586
                 /note="LTR38 repeat: matches 46..611 of consensus"
repeat_region   24777..25093
                 /note="MER6 repeat: matches 2..802 of consensus"
repeat_region   25107..25470
                 /note="AluSx repeat: matches 1..312 of consensus"
repeat_region   25474..27051
                 /note="THE1C repeat: matches 1..370 of consensus"
repeat_region   27052..27421
                 /note="THE1C-internal repeat: matches 5..1580 of
                 consensus"
                 /note="THE1C repeat: matches 1..369 of consensus"

repeat_region   27806..28116
                 /note="AluY repeat: matches 1..309 of consensus"
repeat_region   28722..29399
repeat_region   29400..29708
                 /note="L1MD2 repeat: matches 5302..6111 of consensus"
repeat_region   29709..29861
                 /note="AluSx repeat: matches 1..310 of consensus"
repeat_region   30862..31350
                 /note="L1MD2 repeat: matches 5164..5302 of consensus"
repeat_region   31360..31887
                 /note="HERVL repeat: matches 4611..5143 of consensus"
repeat_region   32357..32689
                 /note="MER31B repeat: matches 2..477 of consensus"
repeat_region   32948..33222
                 /note="LTR16A repeat: matches 91..444 of consensus"
repeat_region   33717..34028
                 /note="AluSx repeat: matches 1..286 of consensus"
repeat_region   34911..35200
                 /note="AluY repeat: matches 1..310 of consensus"
repeat_region   35384..35496
                 /note="AluSp repeat: matches 6..302 of consensus"
repeat_region   36388..36686
                 /note="MIR repeat: matches 32..147 of consensus"
repeat_region   37165..37220
                 /note="AluSx repeat: matches 1..302 of consensus"
repeat_region   37227..37266
                 /note="7 copies 8 mer tccttcct 75% conserved"
repeat_region   38122..38421
                 /note="MIR repeat: matches 211..250 of consensus"
repeat_region   38644..38737
                 /note="AluSx repeat: matches 11..310 of consensus"
repeat_region   39325..39632
                 /note="L2 repeat: matches 2616..2717 of consensus"
repeat_region   40283..40636
                 /note="AluJo repeat: matches 4..310 of consensus"
                 /note="THE1A repeat: matches 1..353 of consensus"

Query Match   69.3%;   Score 578.4;   DB 9;   Length 100296;
Best Local Similarity 99.8%;   Pred. No. 5.2e-121;
Matches 579; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy   256 AAATGCTGATGGAAAGTTAAGCTGAAATTTGGGGTCCTCTTCCGTGATGATAAATGTGC 315
           | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   74867 AGATGCTGATGGAAAGTTAAGCTGAAATTTGGGGTCCTCTTCCGTGATGATAAATGTGC 74926

Qy   316 CAACCTCTTGAAGCAGAGCTGCTTCTCGCAAGTGTTCATGATGATGTTGACATTATTACTGCA 375
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   74927 CAACCTCTTGAAGCAGAGCTGCTTCTCGCAAGTGTTCATGATGATGTTGACATTATTACTGCA 74986

Qy   376 ATATCCAGGAGAGCTGCTTCTCGCAAGTGTTCATGATGATGTTGACATTATTACTGCA 435
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   74987 ATATCCAGGAGAGCTGCTTCTCGCAAGTGTTCATGATGATGTTGACATTATTACTGCA 75046

Qy   436 AGATTAAATGTGGTTTACATATCTTATGTACTGCCATTTTTGTTTCTGGTAAACTGGAA 495
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   75047 AGATTAAATGTGGTTTACATATCTTATGTACTGCCATTTTTGTTTCTGGTAAACTGGAA 75106

Qy   496 TATAAAGTGAAAGAACAAACATTTGAACATACTTAAATGTATTTTTATATCTTGTAAA 555
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   75107 TATAAAGTGAAAGAACAAACATTTGAACATACTTAAATGTATTTTTATATCTTGTAAA 75166

Qy   556 CGAAAGGAGGATTCATGTTTTAGAAGTCTGTCGTCCTTTTTATATCTTGAAAGAAAATCTATG 615
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   75167 CGAAAGGAGGATTCATGTTTTAGAAGTCTGTCGTCCTTTTTATATCTTGAAAGAAAATCTATG 75226

Qy   616 TATGATGCTATAAAATAAAATCCTATTATTTTCTCAGGAATCCTGGTTAGGAATTGCAGGC 675
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   75227 TATGATGCTATAAAATAAAATCCTATTATTTTCTCAGGAATCCTGGTTAGGAATTGCAGGC 75286

Qy   676 AATGAGATTTTTGCGGGGCAGGGATGGGAAATGTTTGTTCATAAATAATTAGACATTTTC 735
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   75287 AATGAGATTTTTGCGGGGCAGGGATGGGAAATGTTTGTTCATAAATAATTAGACATTTTC 75346

us-09-648-310-3.rge

Thu Mar 18 09:42:43 2004

```
    830 TGTCTT  835
        ||||||
     12 TGTCTT   7

RESULT 8
AX408040/c
LOCUS       AX408040     406 bp    DNA    linear   PAT 14-JUN-2002
DEFINITION  Sequence 687 from Patent WO0229103.
ACCESSION   AX408040
VERSION     AX408040.1  GI:21440745
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Alvares,C., Horne,D., Peres-da-Silva,S. and Vockley,J.G.
  TITLE     Gene expression profiles in liver cancer
  JOURNAL   Patent: WO 0229103-A 687 11-APR-2002;
            GENE LOGIC INC (US)
FEATURES             Location/Qualifiers
     source          1..406
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="EMBL/GenBank Accession No. AA287347"
ORIGIN

Query Match             45.8%;  Score 382.4;  DB 6;  Length 406;
Best Local Similarity   99.7%;  Pred. No. 1.6e-76;
Matches   383;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy    452 CATATCCTTATGTACTGCCATTTTTTGTTCTCGGTAAACTGGAATATAAAGTGAAAGAAC 511
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    406 CATATCCTTATGTACTGCCATTTTTTGTTCTCGGTAAACTGGAATATAAAGTGAAAGAAC 347

Qy    512 AAACATTTGAACATACTTAATGTATTTTTATAGAACTTTGTAAACGAAGGAGATTCATG 571
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    346 AAACATTTGAACATACTTAATGTATTTTTATAGAACTTTGTAAACGAAGGAGATTCATG 287

Qy    572 TTTTAGAAGTCTGTCCTTTTTATATCTTGAAAGAAAATCATGTGATGCTATAAAAT 631
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    286 TTTTAGAAGTCTGTCCTTTTTATATCTTGAAAGAAAATCATGTGATGCTATAAAAT 227

Qy    632 AAATCCTATTATTTTTCTCAGGAATCGGTTAGGAATGCAGGCAATGAGATTTTTGCG 691
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    226 AAATCCTATTATTTTTCTCAGGAATCGGTTAGGAATGCAGGCAATGAGATTTTTGCG 167

Qy    692 GGGCAGGGATGGGAATGTTTGTTGTTCATAAATAATTAGACATTTTCTATAGATATTGACAT 751
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    166 GGGCAGGGATGGGAATGTTTGTTGTTCATAAATAATTAGACATTTTCTATAGATATTGACAT 107

Qy    752 TCTGCGAAAGCAACAAGCAAACTGAAGACCAACTCCTATGAGAAATATTATGATGTTTAT 811
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    106 TCTGCGAAAGCAACAAGCAAACTGAAGACCAACTCCTATGAGAAATATTATGATGTTTAT 47

Qy    812 GTAATAAAGACATGTAACTGTCTT 835
          ||||||||||||||||||||||||
Db     46 GTAATAAAGACATGTAACTGTCTT 23

RESULT 9
AX341854/c
LOCUS       AX341854     368 bp    DNA    linear   PAT 10-JAN-2002
DEFINITION  Sequence 2101 from Patent WO0196388.
ACCESSION   AX341854
VERSION     AX341854.1  GI:18137836
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
```

```
    736 TATAGATATTTGACATTCGCGAAAGCAACAAGCAAACTGAAGACCAACTCCTATGAGAA 795
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
  75347 TATAGATATTTGACATTCGCGAAAGCAACAAGCAAACTGAAGACCAACTCCTATGAGAA 75406

    796 ATATTATGATGTTTATGTAATAAAGACATGTAACTGTCTT 835
        ||||||||||||||||||||||||||||||||||||||||
  75407 ATATTATGATGTTTATGTAATAAAGACATGTAACTGTCTT 75446

AX396724/c
LOCUS       AX396724     552 bp    DNA    linear   PAT 18-MAY-2002
DEFINITION  Sequence 939 from Patent WO0212328.
ACCESSION   AX396724
VERSION     AX396724.1  GI:21067471
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   King,G.E., Meagher,M.J., Xu,J. and Secrist,H.
  TITLE     Compositions and methods for the therapy and diagnosis of colon
            cancer
  JOURNAL   Patent: WO 0212328-A 939 14-FEB-2002;
            CORIXA CORPORATION (US)
FEATURES             Location/Qualifiers
     source          1..552
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

Query Match             65.0%;  Score 542.8;  DB 6;  Length 552;
Best Local Similarity   99.6%;  Pred. No. 4.9e-113;
Matches   544;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy    290 GTCCTCCTTCCGTGATGATAAAATGTGCCAACTCTTGAAGCATTGTAAGGTGTTAAA 349
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    552 GTCCTCCTTCCGTGATGATAAAATGTGCCAACTCTTGAAGCATTGTAAGGTGTTAAA 493

Qy    350 GCTGCAAAACGAAGGAAGAGATTGTAACATATCCAGGAGAGCTGCTTCTGCAAGGTGTTCAT 409
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    492 GCTGCAAAACGAAGGAAGAGATTGTAACATATCCAGGAGAGCTGCTTCTGCAAGGTGTTCAT 433

Qy    410 GATGATGTTGACATTATATTACTCGCAAGATAATGTGGTTTACATATCTTATGTACTGC 469
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    432 GATGATGTTGACATTATATTACTCGCAAGATTAATGTGGTTTACATATCTTATGTACTGC 373

Qy    470 CATTTTTTGTTTCTGGTAAACTGGAAATTAAAGTGAAGAACAAACATTTGAACATACTT 529
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    372 CATTTTTTGTTTCTGGTAAACTGGAAATTAAAGTGAAGAACAAACATT TGAACATACTT 313

Qy    530 AATGTATTTTTATAGAACTTTGTAAACGAAAGGAGATTCATGTTTTAGAAGTCTGTCCTT 589
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    312 AATGTATTTTTATAGAACTTTGTAAACGAAAGGAGATTCATGTTTTAGAAGTCTGTCCTT 253

Qy    590 TTTTATATCTTGAAAGAAAATCATGTGATGCTATAAAATAAATCCTATTATTTTTTCT 649
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    252 TTTTATATCTTGAAAGAAAATCATGTGATGCTATAAAATAAATCCTATTATTTTTTCT 193

Qy    650 CAGGAATCTCGTTAGGAATTGCAGGCAATGCAGGCAATGGGGGCAGGGATGGGAATGT 709
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    192 CAGGAATCTCGTTAGGAATTGCAGGCAATGCAGGCAATGGGGGCAGGGATGGGAATGT 133

Qy    710 TTGTTCATAAATAATTAGACATTTTCTATAGATATTGACATTTCTATAGATATTTGACATTCGCGAAAGCAACAAGC 769
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    132 TTGTTCATAAATAATTAATTAGACATTTTCTATAGATATTTGACATTCGCGAAAGCAACAAGC 73

Qy    770 AAACTGAAGACCAACTCCTATGAGAAATATTATGATGTTTATGTAATAAAGACATGTAAC 829
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     72 AAACTGAAGACCAACTCCTATGAGAAATATTATGATGTTTATGTAATAAAGACATGTAAC 13
```

REFERENCE   1
AUTHORS   Jiang,Y., Harlocker,S.L. and Secrist,H.
TITLE   Compositions and methods for the therapy and diagnosis of colon cancer
JOURNAL   Patent: WO 0196388-A 2101 20-DEC-2001;
        CORIXA CORPORATION (US)
FEATURES   Location/Qualifiers
source   1..368
        /organism="Homo sapiens"
        /mol_type="unassigned DNA"
        /db_xref="taxon:9606"
ORIGIN

Query Match   43.7%;   Score 365.2;   DB 6;   Length 368;
Best Local Similarity   99.2%;   Pred. No. 1.3e-72;
Matches   365;   Conservative   1;   Mismatches   2;   Indels   0;   Gaps   0;

```
Qy     43 CCTGCTCCTCCCCACAGCCGGTCTTCTTCTTTGCCTCAGCCACTTCCTTCCTTCGCC 102
          ||||||||||||||||   |||| |||||||  ||||||||||||||||||||||||
Db     16 CGTCCCCTCCCTTCCCACTGCAGCCTTTCTTCTTAGCCCGGAACCACTTCCTTCTTCGCT 75

Qy    103 TCACCCTCCCAGTGCACTGCAGAAGAGTAACCGGGTCAGACCCAGCGCGGCGCCAGTTCT 162
          |||||||||||||||| |||||
Db     76 TGTTCCTCCCTAGGGCGCGGAAGCTGAGTCGAGGGTTCAGACCCCAGCGGGCGAGCAGCTC 135

Qy    163 CCGGCGGAAGGAAAACCGGCGCCAGAGAGGCCAGAGAGCAATGAATGATTGTGATCACGAGGTTAACCT 222
          ||   ||||| |||| |||||  ||||||||||||||||||||||||
Db    136 TTCAGTGAAGAAGGAAGGAAGCAATCGGAGGGTCAGCAATGAACGTGAGCATGAGGTTAACCT 195

Qy    223 CTTAGTGGAGGAAATTCATCGTTCTGGGTTCAAAAAATGCTGATGGAAAGTTAAGCGTGAA 282
          |||||| |||||||||||||||||  |||||||||| || |  |  | ||||
Db    196 CCTGGGTGGAGGAAATTCATCGTCTCTGGGTTCCAAAAAATGCCGATGGGAAACTGAGTGTGAA 255

Qy    283 ATTTGGGGTCCTCTTCCGTGATGATAAAGTGCCAACCTCTTTGAAGCATTGGTAGGAAC 342
          |||| || ||||||||| ||  |||  | |||  ||   ||   |||
Db    256 GTTTGGGGTCCTCTTCCAAGACGACGACGATGTGCCAATCTCTTTGAAGCGTTGGTGGGAAC 315

Qy    343 TCTTAAAGCTGCAAAACGAAGGAAGATTGTACATATCCAGGAGGAGACGCTGCTTCTGCAAGG 402
          ||| || |||||||||||||||||||||||  |||||||||||||||||||||||||
Db    316 TCTGAAAGCCGCAAAACGAACGAAGGAAGATTGTACGTACGCAGGAGGAGCTGCTTTGCAAGG 375

Qy    403 TGTTCATGATGGTTGACATTATTATTACTGCAAGATTAATGTGGTTTACATATCTTTAT 462
          ||||||||||||| ||||||||||  |||||||||||||||||||||||||
Db    376 TGTTCATGATGTGTGACATTGTTGACATTGCTGCAAGATTAATGTGGTTGCAGAATCGTGGGG 435

Qy    463 GTACTGCCATTTTTGTTTCTGGTAAACTGGAAT-ATAAGTGAAAGAACAAACATTTGA 521
          |||||||||||||        |||||||||  |||||||||||||||||||||||||
Db    436 GTA-----------TCTGGTAAACTGGAATAATTAAGTTAAGGACAAACAT---G 477

Qy    522 ACATACTTAATGTATTTTATAGAACTTGTAAAGGAGATTCATGTTTTAGAAGT 581
          ||||||||||||||||||||||||||  |||||||||||||||||
Db    478 AAGTTCCTTATGTATTTTATAGACCTTTGTAAACAAAGGGGA--CTTGTTGAGAGTC 535

Qy    582 CTGTCCTTTTTATATCTTGAGAAGAAAATCTATGTATGATGCTATAAAATAAATCCTATT 641
          ||||| ||||||||||||  ||||  ||||||||||||||||||||||||||
Db    536 CTGTTTTTTATACCTTGGAGCAAAAACATTACAATGTAAAATAAACAAACCTGTTTATTT 595

Qy    642 ATTTTTCTCAGGAATCTGGTTAGGAATTGCAGGCAATGAGAGATTTTTTGCGGGGCAGGGAT 701
          ||||||||||||||||||  || ||||  ||||||  |||||
Db    596 TTTTTTCTTAAGGAGGTAATCGGGAGGACGTAGGCCAATAAAATGTTTTCAGAGGTGCGAAA 655

Qy    702 GGGAATGTTGTTCATAAATAATTAGACATTTCTATAGATATTTGACATTCTGCGAAAG 761
          ||||||  ||  |||| | || |||  ||
Db    656 AAGCTTTTGTTTTCTTAAACCATTCTTAGTCTT--CTGCCACACTTGACACTCCGTCAAAG 713

Qy    762 CAACAAGCAAACTGAAGACCAACTCCTATGAGAAATATTATGATGTTTATGTAATAAAGA 821
          ||| || ||| | |||||||||||||  ||||| ||||||||
Db    714 TCAGAAGCGAACTAAGACAAGCAACTGCGGTGAAAATATTATTATTGTTTATGTAATAAAAAAA 773

Qy    822 CATGTAAC 829
          ||||||||
Db      8 CATGTAAC 1
```

RESULT 11
AX456994
LOCUS   AX456994   780 bp   DNA   linear   PAT 06-JUL-2002
DEFINITION   Sequence 5 from Patent WO0216419.
ACCESSION   AX456994
VERSION   AX456994.1   GI:21715785
KEYWORDS   .
SOURCE   Rattus sp.
ORGANISM   Rattus sp.
        Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
        Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
        Rattus.
REFERENCE   1
AUTHORS   Fisher,P.B., Kang,D.C. and Su,Z.Z.
TITLE   Progression suppressed gene 13 (psgen 13) and uses thereof
JOURNAL   Patent: WO 0216419-A 5 28-FEB-2002;
        The Trustees of Columbia University in the City of New York (US)
FEATURES   Location/Qualifiers
source   1..780

RESULT 10
AX456990
LOCUS   AX456990   780 bp   DNA   linear   PAT 06-JUL-2002
DEFINITION   Sequence 1 from Patent WO0216419.
ACCESSION   AX456990
VERSION   AX456990.1   GI:21715783
KEYWORDS   .
SOURCE   Rattus sp.
ORGANISM   Rattus sp.
        Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
        Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
        Rattus.
REFERENCE   1
AUTHORS   Fisher,P.B., Kang,D.C. and Su,Z.Z.
TITLE   Progression suppressed gene 13 (psgen 13) and uses thereof
JOURNAL   Patent: WO 0216419-A 1 28-FEB-2002;
        The Trustees of Columbia University in the City of New York (US)
FEATURES   Location/Qualifiers
source   1..780
        /organism="Rattus sp."
        /mol_type="unassigned DNA"
        /db_xref="taxon:10118"
ORIGIN

Query Match   40.1%;   Score 335.2;   DB 6;   Length 780;
Best Local Similarity   69.1%;   Pred. No. 9.4e-66;
Matches   539;   Conservative   0;   Mismatches   218;   Indels   23;   Gaps   5;

```
REFERENCE    1  (bases 1 to 692)
AUTHORS      Ganguly,K., Yang,L.F. and Reddy,P.K.
TITLE        A novel cDNA clone from mouse thymus cDNA library
JOURNAL      Unpublished
REFERENCE    2  (bases 1 to 692)
AUTHORS      Ganguly,K., Yang,L.F. and Reddy,P.K.
TITLE        Direct Submission
JOURNAL      Submitted (15-MAY-1998) Anatomy & Cell Biology, SUNY HSC at
             Brooklyn, 450 Clarkson Avenue, Brooklyn, NY 11203, USA
FEATURES             Location/Qualifiers
     source          1..692
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /db_xref="taxon:10090"
                     /tissue_type="thymus"
     CDS             <1..240
                     /codon_start=1
                     /product="unknown"
                     /protein_id="AAL16773.1"
                     /db_xref="GI:16303307"
                     /translation="BFHFFYSINNIFMAAGLQPASCFDRMSSGCRNCLKNGLREQKL
                     FHTSKNVSLPTFPNYLLKKNRFCLFYIVMFCFKV"

ORIGIN

Query Match        37.5%;   Score 312.8;   DB 10;   Length 692;
Best Local Similarity  71.2%;   Pred. No. 1.2e-60;
Matches 494;  Conservative  0;  Mismatches 172;  Indels 28;  Gaps 5;

Qy    132  ACCGGGTCCAGACCACCACGCGGCCGCCCAGTTCTCCGGCGGGAAGGAAAACCGCGCAGAGAGG  191
           ||||||||| | |||  | || |  | ||||||| ||||||| ||||||  |||| ||||||||
Db    678  ACAGGGTTCAGACCCCACCACGCAGCAGCAGCTCTTCCGTGAAGAAG-GAAGCAATCGGAGAGT  620

Qy    192  CAGCAATGAAATGTGGATCACGAGGTTAACCTCTTAGTGGAGGAAATTCATCGTTTGGGTT  251
           ||||||||||||||||||| | |||||||||||||||||||||||||||||||| ||||||
Db    619  CAGCAATGAAATGTGGAACATGAAGTTAACCTCCTGGTGGAGGAAATTCATCGCCTGGGTT  560

Qy    252  CAAAAAAATGCTGATGAGAAATTGGTGAAATTAAGTGTGAGGAACTCTTAAGCTGCAAAA  311
           |||||| ||||| |||||||||||||| |||||||||||| ||||| |||||||| ||||
Db    559  CCAAAAATGCTGATGAAAAAATTAAGTGTGAAAAATTAAGTGTGAAGGTTGGGGCCTCTTCCAGGATGACAGAT  500

Qy    312  GTGCCAACCTCTTTGAAGCATTGGTAGGAACTCTAAAGCTGCAAAACGAAGGAAGATTG  371
           ||||||||||||| |||||||||||| |||| |||||||||||||||| ||||| |||
Db    499  GTGCCAATCTCTTTGAAGCGTTGGTAGGAACTCTGAAAGCTGCAAAACGAAGGAAGAGTG  440

Qy    372  TAACATATCCAGGAGAGCTGCTTCTCGCAAGGTGTTCATGATGATGTGACATTATATTAC  431
           |||| ||| |||||||||| | |||||||||| ||| |||||||| |||||| ||||||
Db    439  TTACATACGCAGGGGAACTACTTTTGCAAGGTGTTCAAGGTGTTCAAGTTGACATTGTATTGC  380

Qy    432  TGCAAGGATTAATGTGGTTTACATATCTTTATGTACTGCCATTTTTTGTTCTGGTAAACT  491
           |||||| |||||||||||||| |||| ||||| ||| |||||| |||
Db    379  TGCAAGATTAATGTGGTTTGCATGGCTTGGTGTA-------------TCTGATAAACT  335

Qy    492  GGAATAAAGTGAAGCAATGAAACAATTGAACATTTTATAGTATTTTATAGAACTTTG  551
           |||| ||||  ||||||| ||| |||||||| |||| ||||||| ||||||||||||
Db    334  GGAATAACTAAGTAAAGACTAACGTGAATTTCCTTATGTGTATATTTTATAGAACTTTG  275

Qy    552  TAAACGAAAGGAGATTCATGTTTAGAAGTCTGGTCCTTTTTTTATATCTTGAAGAAAATC  611
           ||||| |  ||  ||       | |||| ||||| ||||  |||| ||||| ||| ||
Db    274  TAAACAAAGGGG------GCTTGTTGAGAAGTCCTGTTTTTTATACCTTGAAGCAAAACA  221

Qy    612  TATGTATGATGCTATAAAAATAAATCCTATTATTTTTCTCCAGGAATCTGGTTAGGAATTGC  671
           ||| ||  | ||| |  ||  || |||| ||| |||||||||| |||||||||||||||||
Db    220  TTACAATG--TAAAATAAACAAAACCTATTATTATTTTTTTTACCTTAAGAGGTAATTGGGAAAATGT  163

Qy    672  AGGCAATGAGATGATTTTTCGGGGGACGGATGGAAATGTTTGTTCATAAATAATT----AG  727
           ||||| ||| |  |||| |||||| || ||||  |||| || |||| | || |||    ||
Db    162  AGGTAATGAAACATTTTTGGAGGTGTGAAAAAAGCTTTTTGTTCTCTTAAACCATTCTTAAG  103

Qy    728  ACATTTCTATAGATATTGACATTCGGAAAGCAACAAGCAAACTGAAGACCAACTCC  787
           ||| |||||  || || ||||||| || |||||||||||||| |||||||||||||
Db    102  ACAATTTCTACAGCCACTTGACATTCGTCAAGCAAGCAAGCAAACTGCAGACCAGCTGC  43

Qy    788  TATGAGAAATATTATGATGTTTATGTAATAAAGA  821
```

```
              |||| ||   |  || |    || |   ||| ||||
        42 CATGAAAATGTTGTTGTTTATGGAATAAAAAAAA 9
```

SULT 13
918409
CUS   AX918409   296 bp   DNA   linear   PAT 18-DEC-2003
FINITION Sequence 34272 from Patent EP1033401.
CESSION   AX918409
RSION   AX918409.1 GI:40212198
YWORDS   .
URCE   Homo sapiens (human)
ORGANISM Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
FERENCE   1
AUTHORS   Dumas Milne Edwards,J.B., Duclert,A. and Giordano,J.Y.
TITLE   Expressed sequence tags and encoded human proteins
JOURNAL   Patent: EP 1033401-A 34272 06-SEP-2000;
          Genset (FR)
ATURES   Location/Qualifiers
source   1..296
          /organism="Homo sapiens"
          /mol_type="unassigned DNA"
          /db_xref="taxon:9606"
IGIN

Query Match   33.5%;   Score 279.6; DB 6; Length 296;
Best Local Similarity 99.3%;   Pred. No. 4e-53;
Matches   290; Conservative   1; Mismatches   0; Indels 1; Gaps   1;

    487 AAACTGGAATATAAAGTGAAAGAACAAACATTGAACATACTTAATGTATTTTTATAGAA 546
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
      1 AAACTGGAATATAAAGTGAAAGAACAAACATTGAACATACTTAATGTATTTTTATAGAA 60

    547 CTTTGTAAACGAAAGGAGATTCATGTTTTAGAGTCTGTCCTTTTTATATCTTGAAAGA 606
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
     61 CTTTGTAAACGAAAGGAGATTCATGTTTTAGAGTCTGTCCTTTTTATATCTTGAAAGA 120

    607 AAATCTATGTATGATGCTATAAAATAAATCCTATTATTTTCTCAGGAATCTGGTTAGGA 666
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
    121 AAATCTATGTATGATGCTATAAAATAAATCCTATTATTTTCTCAGGAATCTGGTTAGGA 180

    667 ATTGCAGGCAATGAGATTTTTGCGGGGCAGGGATGGGAATGTTTGTTCATAAATAATTA 726
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
    181 ATTGCAGGCAATGAGATTTTTGCGGGGCAGGGATGGGAATGTTTGTTCATAAATAATTA 240

    727 GACATTTTCTATAGATATTTGACATTCTGCGAAAGCAACAAGCAAACTGAAG 778
        ||||||||||||||||||||||||||||||||| :
    241 GACATTTTCTATAGATATTTGACATTCTGCGAAA-SAACAAGCAAACTGAAG 291
```

BSULT 14
D053942
DCUS   BD053942   296 bp   DNA   linear   PAT 27-AUG-2002
EFINITION Sequence tag and encoded human protein.
CESSION   BD053942
ERSION   BD053942.1 GI:22599548
EYWORDS   JP 2001269182-A/30188 .
URCE   Homo sapiens (human)
ORGANISM Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
EFERENCE 1 (bases 1 to 296)
AUTHORS   Edwards,J.B.D.M., Duclair,E. and Jordan,J.Y.
TITLE   Sequence tag and encoded human protein
JOURNAL   Patent: JP 2001269182-A 30188 02-OCT-2001;
          GENSET
OMMENT   OS   Homo sapiens (human)
          PN   JP 2001269182-A/30188
          PD   02-OCT-2001
          PF   24-FEB-2000 JP 2000118773
          PR   26-FEB-1999 US   60/122487

PI   JEAN BAPUTIST DUMAS MILNE EDWARDS,EIMERIC DUCLAIR,JEAN YVES
PI   JORDAN
PC   C12N15/09,C07K14/435,C07K16/18,C12N1/15,C12N1/19,C12N1/21, PC
      C12N5/10,
PC   C12P21/02,C12P21/08,C12Q1/68//G06F17/30,C12N15/00,C12NS/00, PC
      G06F15/40

CC
FH   Key   Location/Qualifiers.
FEATURES   source
              1..296
              /organism="Homo sapiens"
              /mol_type="genomic DNA"
              /db_xref="taxon:9606"

ORIGIN

Query Match   33.5%;   Score 279.6; DB 6; Length 296;
Best Local Similarity 99.3%;   Pred. No. 4e-53;
Matches   290; Conservative   1; Mismatches   0; Indels 0; Gaps   1;

Qy 487 AAACTGGAATATAAAGTGAAAGAACAAACATTGAACATACTTAATGTATTTTTATAGAA 546
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 AAACTGGAATATAAAGTGAAAGAACAAACATTGAACATACTTAATGTATTTTTATAGAA 60

Qy 547 CTTTGTAAACGAAAGGAGATTCATGTTTTAGAGTCTGTCCTTTTTATATCTTGAAAGA 606
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 CTTTGTAAACGAAAGGAGATTCATGTTTTAGAGTCTGTCCTTTTTATATCTTGAAAGA 120

Qy 607 AAATCTATGTATGATGCTATAAAATAAATCCTATTATTTTCTCAGGAATCTGGTTAGGA 666
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 121 AAATCTATGTATGATGCTATAAAATAAATCCTATTATTTTCTCAGGAATCTGGTTAGGA 180

Qy 667 ATTGCAGGCAATGAGATTTTTGCGGGGCAGGGATGGGAATGTTTGTTCATAAATAATTA 726
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 181 ATTGCAGGCAATGAGATTTTTGCGGGGCAGGGATGGGAATGTTTGTTCATAAATAATTA 240

Qy 727 GACATTTTCTATAGATATTTGACATTCTGCGAAAGCAACAAGCAAACTGAAG 778
        ||||||||||||||||||||||||||||||||| :
Db 241 GACATTTTCTATAGATATTTGACATTCTGCGAAA-SAACAAGCAAACTGAAG 291

RESULT 15
BC039801
LOCUS   BC039801   526 bp   mRNA   linear   ROD 16-APR-2003
DEFINITION Mus musculus RIKEN cDNA 3110003A17 gene, mRNA (cDNA clone
           IMAGE:1448067), partial cds.
ACCESSION BC039801
VERSION   BC039801.1 GI:24980967
KEYWORDS   .
SOURCE   Mus musculus (house mouse)
ORGANISM   Mus musculus
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 526)
AUTHORS   Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
           Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
           Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
           Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
           Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
           Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
           Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
           Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
           Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
           McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
           Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
           Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
           Fahey,J., Helton,E., Ketteman,M., Madan,A., Rodrigues,S.,
           Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
           Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
           Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
           Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E.,
           Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
TITLE   Generation and initial analysis of more than 15,000 full-length
           human and mouse cDNA sequences

```
Qy   312 GTGCCAACCTCTTTGAAGCATTGGTAGGAGAACTCTTAAAGTGCAAAACGAAGGAAGATTG 371
             ||||||| ||||||||||||||||  ||||||||||||||| |||||||||||||||||||
Db   187 GTGCCAATCTCTTTGAAGCGTTGGTAGGAGAACTCTGAAAGTCGAAAACGAAGGAAGATTG 246

Qy   372 TAACATATCCAGGAGAGCTGCTTCTGCAAGTGTTCATGATGATGTTGACATTATATTAC 431
             |||||| |||||||||| |||||||||||||||||||||||||| |||| ||||
Db   247 TTACATACGCAGGGGAACTACTTTTGCAAGTGTTCATGATGATGTTGACATTGTATTGC 306

Qy   432 TGCAAGATTAATGTGGTTTACATATCTTTATGTACTGCCATTTTTGTTTCTGGTAAACT 491
             |||||||||||||||||| ||  |||     |||||           ||| ||||||
Db   307 TGCAAGATTAATGTGGTTTGCATGGCTTGGTGTA-------------TCTGATAAACT 351

Qy   492 GGAATATAAAGTGAAAGAACAAACATTTGAACATACTTAATGTATTTTATAGAACTTTG 551
             |||||      |||||||||||||||||||||||||||  ||||||||||||||||||
Db   352 GGAATAACTAAGTTA--AAAGACTAACGTGAATTTCCTTATGTATTTTTAGAACTTTG 409

Qy   552 TAAACGAAAGGAGATTCATGTTTTAGAAGTCTGTCCTTTTTTATATCTTGAAAGAAAATC 611
             ||||   ||||        |||||||||||  || | ||||| |||||||||||
Db   410 TAAACAAGGGGG------GCTTGTTGGAGAAGTCCTGTTTTTTATACCTTGAAGCAAAACA 463

Qy   612 TATGTATGATGCTATAAAATAAATCCTATTATTTTTCTCAGGAA 655
             ||||| ||  |||||||||||||||||||||||||
Db   464 TTACAATG--TAAAATAAACAAAACCTATTATTTTTCTTAAAAA 505
```

Search completed: March 17, 2004, 13:36:35
Job time : 3710 secs

FEATURES   Location/Qualifiers
     source   1..526
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:1448067"
                     /tissue_type="Mammary gland"
                     /clone_lib="Soares_mammary_gland_NbMMG"
                     /lab_host="DH10B"
                     /note="Vector: pT7T3D-Pac"
     gene   <1..526
                     /gene="3110003A17Rik"
                     /note="synonym: 1700028I04Rik"
                     /db_xref="LocusID:73112"
                     /db_xref="MGI:1920362"
     CDS   <1..317
                     /gene="3110003A17Rik"
                     /codon_start=3
                     /product="3110003A17Rik protein"
                     /protein_id="AAH39801.1"
                     /db_xref="GI:24980968"
                     /db_xref="LocusID:73112"
                     /db_xref="MGI:1920362"
                     /translation="RVQGSDPRSSSSSVKKEAIGESAMNVEHEVNLLVEEIHRLGSKN
                     ADGKLSVKFGVLFQDDRCANLFEALVGTLKAAKRRKIVTYAGELLLQGVHDDVDIVLL
                     QD"
ORIGIN

Query Match   29.8%; Score 249.2; DB 10; Length 526;
Best Local Similarity 74.4%; Pred. No. 3.5e-46;
Matches 390; Conservative   0; Mismatches 108; Indels 26; Gaps   5;

```
Qy   132 ACCGGGTCCAGACCCACGCGGCGCCCAGTTCTCCGGCGGGAAGGAGAAAACCGCGCAGAGAGG 191
             ||||||||||| ||||||| ||||||||||| ||||||||||||  ||||  ||| |||||||
Db     8 ACAGGGTTCAGACCCACGCCACGCAGCAGCTCTTCCGTCGAAGAAG-GAAGCAATCGGAGAGT 66

Qy   192 CAGCAATGAATGTGGATCACGAGAGGTTAACCTCTTAGTGGAGGAAAATTCATCGTTTGGGTT 251
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    67 CAGCAATGAATGTGGAACATGAAGTTAACCTCCTGGTGGAGGAAAATTCATCGCCTGGGTT 126

Qy   252 CAAAAAATGCTGATGAAAGTGAAAGTTAAGCGTGAAATTTGGGGTGTCCTCTTCCCGTGATGATAAAT 311
             | ||||||||||||||||||||||||||||||||||| |||||| ||| ||||| |||| | |||
Db   127 CCAAAAATGCTGATGAAAATTAAGTGTGGAAGTTTGGGGTTTGGGTCCTTCTTCCCAGATGACAGAT 186
```

```
            GenCore version 5.1.6
      Copyright (c) 1993 - 2004 Compugen Ltd.

nucleic - nucleic search, using sw model

on:    March 17, 2004, 11:44:37 ; Search time 440 Seconds
                                   (without alignments)
                          8061.921 Million cell updates/sec

le:          US-09-648-310-3
fect score:  835
uence:       1 ggcacgaggcttgagcgcag.........taaagacatgtaactgtctt 835

ring table:  IDENTITY_NUC
             Gapop 10.0 , Gapext 1.0

rched:       3373863 seqs, 2124099041 residues

al number of hits satisfying chosen parameters:    6747726

.imum DB seq length: 0
.imum DB seq length: 2000000000

t-processing: Minimum Match 0%
              Maximum Match 100%
              Listing first 45 summaries

abase :   N_Geneseq_29Jan04:*
          1:  geneseqn1980s:*
          2:  geneseqn1990s:*
          3:  geneseqn2000s:*
          4:  geneseqn2001as:*
          5:  geneseqn2001bs:*
          6:  geneseqn2002s:*
          7:  geneseqn2003as:*
          8:  geneseqn2003bs:*
          9:  geneseqn2003cs:*
         10:  geneseqn2004s:*
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

ult No.	Score	% Query Match	Length	DB	Match	ID	Description		
1	835	100.0	835	6	69	24	2	190	8.3	ABK11086	Abk11086 Human Pro	AAX85691	Aax85691 Novel cDN
2	638.4	76.5	653	6	60	25	6	60	7.2	ABT10814	Abt10814 Human bre	ABN49354	Abn49354 Human spl
3	619.4	74.2	778	3	54.8	26	7	2000	6.6	AAF16110	Aaf16110 Human pro	ADA71938	Ada71938 Rice gene
4	549	65.7	712	3	52.6	27	7	19634	6.6	AAC66413	Aac66413 Human sec	ABZ10015	Abz10015 Haematopo
5	542.8	65.0	552	6	52.2	28	6	19634	6.3	ABK45388	Abk45388 cDNA enco	ABZ10161	Abz10161 Haematopo
6	450.2	53.9	594	6	50.6	29	6	113515	6.3	ABQ57055	Abq57055 Human col	ABL34174	Abl34174 Human imm
7	436.8	52.3	717	6	50	30	6	6073	6.1	ABQ57053	Abq57053 Human col	ABL33542	Abl33542 Human imm
8	382.4	45.8	406	6	50	31	6	50	6.0	ABN94189	Abn94189 Gene #687	ABZ00731	Abz00731 Human leu
9	365.2	43.7	368	6	50	32	6	6020	6.0	ABL38512	Abl38512 Human col	ABL32199	Abl32199 Human imm
10	335.2	40.1	780	6	49.4	33	6	6020	6.0	ABK11085	Abk11085 Rat Progr	ABL49308	Abl49308 Human pol
11	310	37.1	800	3	49.2	34	6	5467	5.9	AAZ21517	Aaz21517 Progressi	ABL34109	Abl34109 Human imm
12	279.6	33.5	296	2	49.2	35	6	14950	5.9	AAC30197	Aac30197 Human sec	ABL33257	Abl33257 Human imm
13	279.4	33.5	627	6	49.2	36	6	5488	5.9	ABQ56862	Abq56862 Human col	ABL33456	Abl33456 Human imm
14	237	28.4	458	8	49.2	37	6	6381	5.9	ACH38863	Ach38863 Human foe	ABL32966	Abl32966 Human imm
15	216.4	25.9	351	6	49.2	38	6	6381	5.9	ABA11066	Aba11066 Human ner	ABL34518	Abl34518 Human met
16	188.4	22.6	388	3	49	39	6	6381	5.9	AAC98308	Aac98308 Human col	ABL70243	Abl70243 Chemicall
17	170.2	20.4	611	3	48.8	40	6	6210	5.9	AAA16371	Aaa16371 Human col	ABK28472	Abk28472 DNA trans
18	109.4	13.1	486	6	48.4	41	6	5728	5.8	ABZ12550	Abz12550 Arabidops	ABL32101	Abl32101 Human imm
19	109.4	13.1	492	9	48.2	42	6	13038	5.8	ADE81847	Ade81847 Arabidops	ABL33275	Abl33275 Human imm
20	109.4	13.1	657	3	48.2	43	6	5937	5.8	AAC53673	Aac53673 Arabidops	ABL34543	Abl34543 Human met
21	98.8	11.8	551	3	48.2	44	6	5937	5.8	AAC53910	Aac53910 Zea mays	ABN80151	Abn80151 Human che
22	69.8	8.4	268	6	48.2	45	6	6618	5.8	ABL72646	Abl72646 Corn tass	ABK31284	Abk31284 Signal tr
23	69	8.3	190	2						AAX85712	Aax85712 Novel cDN		

ALIGNMENTS

```
RESULT 1
ABK11086
ID   ABK11086 standard; cDNA; 835 BP.
XX
AC   ABK11086;
XX
DT   05-JUN-2002  (first entry)
XX
DE   Human Progression Suppressed Gene 13 (HuPSGen 13), cDNA.
XX
KW   Human; Progression Suppressed Gene 13; rPSGen 13; cytostatic; cancer;
KW   blood vessel; nasopharyngeal tumour; thyroid tumour; leukaemia; lymphoma;
KW   breast; lung; prostate; ovary; colon; gene; ss.
XX
OS   Homo sapiens.
XX
FH   Key            Location/Qualifiers
FT   CDS            197..442
FT                  /*tag= a
FT                  /product= "Progression suppressed gene 13 protein"
XX
PN   WO200216419-A2.
XX
PD   28-FEB-2002.
XX
PF   27-AUG-2001; 2001WO-US026795.
XX
PR   25-AUG-2000; 2000US-00648310.
XX
PA   (UYCO ) UNIV COLUMBIA NEW YORK.
XX
PI   Fisher PB, Kang D, Su Z;
XX
DR   WPI; 2002-280914/32.
DR   P-PSDB; AAU76533.
XX
PT   New rat and human Progression Suppressed Gene 13 for preventing the
PT   growth of cancer cells and/or new blood vessels, and for treating
PT   patients suffering from a cancer.
XX
PS   Claim 4; Fig 2; 53pp; English.
XX
CC   The invention relates to novel isolated nucleic acids which encode a rat
CC   or human Progression Suppressed Gene 13 (PSGen 13) protein. The nucleic
CC   acids are useful for preventing the growth of cancer cells and/or new
CC   blood vessels, and for treating patients suffering from a cancer, e.g.
```

nasopharyngeal tumour, thyroid tumour, leukaemia, lymphoma, or cancer of
the breast, lung, prostate, ovary or colon. PSGen 13 may also be used to
suppress the transformed phenotype of a malignant cell. Administration of
PSGen 13 gene or protein may result in a decrease in tumour mass, number
of cancer cells, serum tumour marker, tumour metastasis, vascularisation,
perfusion, or rate of tumour growth, improved clinical symptoms, and/or
increased patient survival. The present sequence represents the coding
sequence of human Progression Suppressed Gene 13 (HuPSGen 13)

Sequence 835 BP; 246 A; 160 C; 176 G; 253 T; 0 U; 0 Other;

Query Match   100.0%;   Score 835; DB 6; Length 835;
Best Local Similarity 100.0%; Pred. No. 9.6e-201;
Matches 835; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
    1 GGCACGAGGCTTGAGCGGCAGAAACACTTACTTTCCCCTACCCTGCTCCTCCTCCA 60
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
    1 GGCACGAGGCTTGAGCGGCAGAAACACTTACTTTCCCCTACCCTGCTCCTCCTCCA 60

   61 CAGCCGTCTTCTCTTTGCCTCAGCCACTTCCTTCTTCCTCCGCTCCACCCTCCCAGTGCAC 120
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
   61 CAGCCGTCTTCTTTGCCTCAGCCACTTCCTTCTTCCTCCGCTCCACCCTCCCAGTGCAC 120

  121 TGAAGAAGGTAACCGGGTCCAGACCCACGCGGCGGCCCAGTTCTCCGGCGGGAAGGAAAACC 180
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
  121 TGAAGAAGGTAACCGGGTCCAGACCCACGCGGCGGCCCAGTTCTCCGGCGGGAAGGAAAACC 180

  181 GCGCAGAGAGGCAGCAATGAATGTGATCACGAGGTTAACCTCTTAGTGGAGGAAATTCA 240
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
  181 GCGCAGAGAGGCAGCAATGAATGTGATCACGAGGTTAACCTCTTAGTGGAGGAAATTCA 240

  241 TCGTTTGGGTTCAAAAAATGCTGATGGAAAGTTAAGCGTGAAATTTGGGGTCCTCTTCCG 300
      |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
  241 TCGTTTGGGTTCAAAAAATGCTGATGGAAAGTTAAGCGTGAAATTGGGGTCCTCTTCCG 300

  301 TGATGATAAAATGTGCCAACCTCTTTGAAGCATTGTAGGAACTCTTAAAGCTGCAAAACG 360
      |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
  301 TGATGATAAAATGTGCCAACCTCTTTGAAGCATTGTAGGAACTCTTAAAGCTGCAAAACG 360

  361 AAGGAAGATTGTAACATATCCAGGAGAGCTGCTTCTGCAAGGTGTTCATGATGATGTTGA 420
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
  361 AAGGAAGATTGTAACATATCCAGGAGAGCTGCTTCTGCAAGGTGTTCATGATGATGTTGA 420

  421 CATTATTACTGCAAGATTAATGTGGTTTACATATCTTATGTTTATGTTACTGCCATTTTTGTT 480
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
  421 CATTATTACTGCAAGATTAATGTGGTTTACATATCTTATGTTTATGTTACTGCCATTTTTGTT 480

  481 TCTGGTAAACTGGAATATAAAGTGAAAGAACAAACATTTGAACATACTACTTAATGTATTTT 540
      |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
  481 TCTGGTAAACTGGAATATAAAGTGAAAGAACAAACATTTGAACATACTACTTAATGTATTTT 540

  541 ATAGAACTTTGTAAACGAAAGGAGATTCATGTTTTAGAAGTCTGTCCTTTTTTATATCTT 600
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
  541 ATAGAACTTTGTAAACGAAAGGAGATTCATGTTTTAGAAGTCTGTCCTTTTTTATATCTT 600

  601 GAAAGAAAATCTATGTATGATGCTATAAAAATAAATCCTATTATTTTTCCAGGAATCTGG 660
      |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
  601 GAAAGAAAATCTATGTATGATGCTATAAAAATAAATCCTATTATTTTTCCAGGAATCTGG 660

  661 TTAGGAATTGCAGGCAATGAATGAGCAGGATGGGAATGTTTGTTCATAAA 720
      ||||||||||||||||||||||||||||||||||||||||||||||||||
  661 TTAGGAATTGCAGGCAATGAATGAGCAGGATGGGAATGTTTGTTCATAAA 720

  721 TAATTAGACATTTCTATAGATATTGACATTCTGCGAAAGCAACAAGCAAACAAACTGAAGAC 780
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
  721 TAATTAGACATTTCTATAGATATTGACATTCTGCGAAAGCAACAAGCAAACAAACTGAAGAC 780

  781 CAACTCCTATGAGAAATATTATGATGTTTATGTAATAAAGACATGTAACTGTCTT 835
      |||||||||||||||||||||||||||||||||||||||||||||||||||||||
  781 CAACTCCTATGAGAAATTATTATGATGTTTATGTAATAAAGACATGTAACTGTCTT 835
```

ABT10814/c
```
ID   ABT10814 standard; cDNA; 653 BP.
XX
AC   ABT10814;
XX
DT   04-DEC-2002 (first entry)
XX
DE   Human breast cancer associated coding sequence SEQ ID NO: 948.
XX
KW   Human; breast specific gene; breast cancer; differential expression;
KW   cytostatic; gene therapy; gene; ss.
XX
OS   Homo sapiens.
XX
XX
PN   WO200259271-A2.
XX
PD   01-AUG-2002.
XX
PF   25-JAN-2002; 2002WO-US002176.
XX
PR   25-JAN-2001; 2001US-0263757P.
PR   25-APR-2001; 2001US-0286090P.
PR   23-MAY-2001; 2001US-0292517P.
XX
PA   (GENE-) GENE LOGIC INC.
XX
PI   Orr MS, Nation M, Diggans JC, Zeng W;
XX
DR   WPI; 2002-674803/72.
XX
PT   Diagnosing breast cancer in a patient comprises detecting the level of
PT   gene expression in cell or tissue samples, where a differential gene
PT   expression is indicative of breast cancer.
XX
PS   Claim 1; SEQ ID NO 948; 260pp + Sequence Listing; English.
XX
CC   The present invention relates to methods of diagnosing breast cancer in a
CC   patient, which comprise detecting the level of expression in a tissue
CC   sample of two or more genes selected from those shown in ABT09867-
CC   ABT11112, where a differential expression of the genes indicates breast
CC   cancer. The methods are useful in diagnosing, treating, detecting the
CC   progression, and in monitoring treatment of breast cancer in patients.
CC   The methods are also useful as a screening tool for agents that modulate
CC   the onset or progression of breast cancer. The breast cancer genes may be
CC   used as diagnostic markers for the prediction or identification of the
CC   malignant state of breast tissue, for confirming the type and progression
CC   of cancer, and for drug screening and assays. The present sequence is a
CC   coding sequence of the invention. Note: The sequence data for this patent
CC   did not form part of the printed specification, but was obtained in
CC   electronic format directly from WIPO at
CC   ftp.wipo.int/pub.published_pct_sequences
XX
SQ   Sequence 653 BP; 213 A; 135 C; 88 G; 217 T; 0 U; 0 Other;
```

Query Match   76.5%;   Score 638.4;   DB 6;   Length 653;
Best Local Similarity 99.1%;   Pred. No. 3.9e-151;
Matches 642;   Conservative 0; Mismatches 6; Indels 0; Gaps 0;

```
Qy  188 GAGGCAGCAATGAATGTGATCACGAGGTTAACCTCTTAGTGGAGGAAATTCATCGTTTG 247
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  653 GAGGCAGCAATGAATGTGATCACGAGGTTAACCTCTTAGTGGAGGAAATTCATCGTTTG 594

Qy  248 GGTTCAAAAAATGCTGATGGAAAGTTAAGCGTGAAAATTTGGGGTCCTCTTCCGTGATGAT 307
        ||||||||||||||||||||||||||||||||||||| |||||| ||||||||||||||||
Db  593 GGTTCAAAAAATGCTGATGGAAAGTTAAGCGTGAAAATTGGGGGTCCTCTTCCGTGATGAT 534

Qy  308 AAATGTGCCAACCTCTTTGAAGCATTGGTAGGAACTCTTAAAGCTGCAAAACGAAGGAAG 367
        ||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||
Db  533 AAATGTGCCAACCTCTTTGAAGCATTGGTAGGAACTCTTAAAGCTGCAAAAACGAAGGAAG 474

Qy  368 ATTGTAACATATCCAGGAGAGCTGCTTCTGCAAGGTGTTCATGATGATGTTGACATTATA 427
        |||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||
Db  473 ATTGTAACATATCCAGGAGAGCTGCTTCTGCAAGGTGTTCATGATGATGTTGACATTATA 414
```

ESULT 2

The prostate cancer antigens can have neuroprotective, cytostatic, cardioactive, immunomodulatory, muscular, vulnerary, gastrointestinal, nephrotropic, antiinfective, gynaecological and antibacterial activities, and can be used in gene therapy. The prostate cancer antigen polynucleotides may be used for detection of prostate cancer, identification, as chromosome markers, and for numerous other diagnostic or research purposes. The prostate cancer antigens may be used to treat disorders such as neural, immune, muscular, reproductive, gastrointestinal, pulmonary, cardiovascular, renal, and proliferative disorders, wounds, and infectious diseases. AAF16506 to AAF16514 to AAB57303 represent sequences used in the exemplification of the present invention

Sequence 778 BP; 250 A; 151 C; 148 G; 221 T; 0 U; 8 Other;

Query Match   74.2%; Score 619.4; DB 3; Length 778;
Best Local Similarity 99.8%; Pred. No. 2.7e-146;
Matches 620; Conservative 0; Mismatches 0; Indels 1; Gaps 0;

```
Qy    26 CTTACTTTCCCCCTACCCTGCTCCTCCTCTCCAGCGCCGTCTTTCTCTTTGCCTCAGC    85
            |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db    15 CGTACTTTCCCCCTACCCTGCTCCTCCTCTCCCACAGCGCCGTCTTTCTCTTTGCCTCAGC 74

Qy    86 CACTTCCTTCCTTCGCCTCACCCTCCCCAGTGCACTGAAGAGAGGTAACCGGGTCCAGACC  145
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    75 CACTTCCTTCCTTCGCCTCACCCTCCCCAGTGCACTGAAGAGAGGTAACCGGGTCCAGACC  134

Qy   146 CACGCGGCGCCAGTTCTCCGGCGGGAAGGAAAACCGCGCAGAGAGGCAGCAATGATGTG  205
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   135 CACGCGGCGCCAGTTCTCCGGCGGGAAGGAAAACCGCGCAGAGAGGCAGCAATGAATGTG 194

Qy   206 GATCACGAGGTAACCTCTTAGTGAGGAGGAAATTCATCGTTTGGGTTCAAAAAATGCTGAT  265
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   195 GATCACGAGGTAACCTCTTAGTGAGGAGGAAATTCATCGTTTGGGTTCAAAAAATGCTGAT  254

Qy   266 GGAAAGTTAAGCGTGAAATTGGGGTCCTCTTCCGTGATGATAAATGTGCCAACCTCTTT  325
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   255 GGAAAGTTAAGCGTGAAATTGGGGTCCTCTTCCGTGATGATAAATGTGCCAACCTCTTT  314

Qy   326 GAGAGCATTGGTAGGAACTCTTAAAGCTGCAAAACGAAGGAAGATTGTAACATATCCAGGA  385
           |||| ||||||||||||||||||||||| |||||||||||||||||||||||||||||||
Db   315 GAAGCATTGGTAGGAACTCTTAAAGCTCCAAAACGAAGGAAGATTGTAACATATCCAGGA  374

Qy   386 GAGCTGCTTCTGCAAGGTGTTCATGATGAGTGTGACATTAATACTGCAAGATTAATGT  445
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   375 GAGCTGCTTCTGCAAGGTGTTCATGATGAGTGTGACATTAATACTGCAAGATTAATGT  434

Qy   446 GGTTACATATCTTATGTACTGCCATTTTTGTTTCTGGTAAACTGGAATATAAAGTGA  505
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   435 GGTTACATATCTTATGTACTGCCATTTTTGTTTCTGGTAAACTGGAATATAAAGTGA  494

Qy   506 AAGAACAAACATTTGAACATACTTAAGTGTATTTTATAGAACTTGTAAACGAAAGGAGA  565
           ||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db   495 AAGAACAAACATTTGAACATACTTAAGTATTTTTATAGAACTTGTAAACGAAAGGAGA  554

Qy   566 TTCATGTTTTAGAAGTCTGTCCTTTTTTATATCTTGAAAGAAAATCTATGTATGATGCTA  625
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   555 TTCATGTTTTAGAAGTCTGTCCTTTTTTATATCTTGAAAGAAAATCTATGTATGATGCTA  614

Qy   626 TAAAATAAATCCTATTATTTT  646
           |||||||||||||||||||||
Db   615 TAAAATAAATCCTATTATTTT  635
```

RESULT 4
AAC66413
AAC66413
ID   AAC66413 standard; DNA; 712 BP.
XX
AC   AAC66413;
XX
DT   14-FEB-2001 (first entry)
XX
DE   Human secreted protein coding sequence SEQ ID NO: 14.

```
    428 TTACTGCAAGATTAATGTGGTTACATATCTTATGTACTGCCATTTTTGTTTCTGTGTA  487
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
    413 TTACTGCAAGATTAATGTGGTTACATATCTTATGTACTGCCATTTTTGTTTCTGTGTA  354

    488 AACTGGAATATAAAGTGAAAGAACAAACATTTGAACATACTTAAATGTATTTTATAGAAC  547
          |||||||||||||||||||||||||||||||||||||||||||| |||||||||||||||
    353 AACTGGAATATAAAGTGAAAGAACAAACATTTGAACATACTTAATGTATTTTTATAGAAC  294

    548 TTTGTAAACGAAAGGAGATTCATGTTTTAGAAGTCTGTCCTTTTTTATATCTTGAAAGAA  607
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
    293 TTTGTAAACGAAAGGAGATTCATGTTTTAGAAGTCTGTCCTTTTTTATATCTTGAAAGAA  234

    608 AATCTATGTATGATGCTATAAAAATAAATCCTATTATTTTCTCCAGGAATCTGGTTAGGAA  667
          ||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
    233 AATCTATGTATGATGCTATAAAAATAAATCCTATTATTTTCTCAGGAATCTGGTTAGGAA  174

    668 TTGCAGGCAATGACAGATTTTTGCGGGGCCAGGGATGGGAATGTTTGTTGTTCATAAATAATTAG  727
          |||||||||||| |||||||||||||||||||||||||||||||||||||||||||||||||
    173 TTGCAGGCAATGAGATTTTTGCGGGGCCAGGGATGGGAATGTTTGTTGTTCATAAATAATTAG  114

    728 ACATTTCTATAGATATTTGACATTCTGCGAAAGCAACAAGCAAACTGAAGACCAACTCC  787
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
    113 ACATTTCTATAGATATTTGACATTCTGCGAAAGCAACAAGCAAACTGAAGACCAACTCC  54

    788 TATGAGAAATATTATGATGTTTATGTAATAAAGACATGTAACTGTCTT  835
          ||||||||| ||||||||||||||||||||||||||||||||||||||
     53 TATGAGAAAATATTATGATGTTTATGTAATAAAGACATGTAACTGTCTT  6
```

SULT 3
16110
AAF16110 standard; cDNA; 778 BP.
AAF16110;
13-MAR-2001 (first entry)
Human prostate cancer antigen nucleotide sequence SEQ ID NO:545.
Human; prostate cancer; prostate cancer antigen; detection; diagnosis; neuroprotective; cytostatic; cardioactive; immunomodulatory; muscular; vulnerary; gastrointestinal; nephrotropic; antiinfective; gynaecological; antibacterial; gene therapy; neural; immune; reproductive; renal; gastrointestinal; pulmonary; cardiovascular; proliferative disorder; wound; infectious disease; ss.
Homo sapiens.
WO200055174-A1.
21-SEP-2000.
08-MAR-2000; 2000WO-US005988.
12-MAR-1999; 99US-0124270P.
(HUMA-) HUMAN GENOME SCI INC.
(ROSE/) ROSEN C A.
Rosen CA, Ruben SM;
WPI; 2000-587513/55.
P-PSDB; AAB56907.
Prostate cancer associated gene sequences, referred to as prostate cancer antigens, useful for treatment, prevention, and diagnosis of disorders such as prostate cancer.
Claim 1; Page 1004-1005; 2338pp; English.
AAF15566 to AAF16505 encode the human prostate cancer associated proteins, called prostate cancer antigens, given in AAB56363 to AAB57302.

```
Qy    337  AGGAACTCTTAAAGCTGCAAAACGAGGAGGAGAAGATTGTAACATATCCAGGAGAGCTGCTTCT  396
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    269  AGGAACTCTTAAAGCTCGCAAAACGAGGAAGATTGTAACATATCCAGGAGAGCTGCTTCT      328

Qy    397  GCAAGGTGTTCATGATGATGTTGACATTATATTACTGCAAGATTAATATGTCGTTACATAT   456
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    329  GCAAGGTGTTCATGATGATGTTGACATTATTACTGCAAGATTAATATGTCGTTACATAT      388

Qy    457  CTTTAGTGACTGCCATTTTTGTTCTCGGTAAACTGGAATATAAAGTGAAAGAACAAAACA   516
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    389  CTTTATGTACTGCCATTTTTGTTCTCGGTAAACTGGAATATAAAGTGAAAGAACAAAACA   448

Qy    517  TTTGAACATACTTAATGTATTTTATAGAACTTGTGAAAACGAAAAGAGATTCATGTTTTA   576
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    449  TTTGAACATACTTAATGTATTTTATAGAACTTGTAAACGAAAAGAGATTCATGTTTTA    508

Qy    577  GAAGTCTGTCCTTTTTTATATCTTGAAAGAAAATCATGTGATGATGCTATAAAATAAATC   636
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    509  GAAGTCTGTCCTTTTTATATCTTGAAAGAAAATCATGTGATGATGCTATA------      559

Qy    637  CTATTATTTTTCTCAGGAATCTGGTTAGGATCTGGCAATGAGATTTTTTGCGGGGCA     696
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    560  C-----------------------------------------------------      559

Qy    697  GGGATGGGAAATGTTGTTGTTCATAAATAATTAGACATTTCTATAGATATTTGACATTCTGC   756
           ||||||||||||||||||||||||||||||||||||||||||||||
Db    560  ---------------AAATAATTAGACATTTCTATAGATATTTGACATTCTGC      598

Qy    757  GAAAGCAACAAGCAAACTGAAGACCAACTCCTATAGAAAATATTATGATGTTTATGTAAT   816
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    599  GAAAGCACAAGCAAACTGAAGACCAACTCCTATGAGAAATATTATGATGTTTATGTAAT   658

Qy    817  AAAGACATGTAACTGTCTT   835
           ||||||||||||||||||
Db    659  AAAGACATGTAACTGTCTT   677
```

RESULT 5
ABK45388/c

The invention relates to the isolation of genes AAC66410-C66458 encoding
49 human secreted proteins AAB32002-B32050. The genes can be used to
generate fusion proteins by linking to the gene for the human
immunoglobulin G Fc portion (SEQID1) for increasing the stability of the
fusion protein as compared to the human protein only. The genes and
proteins are useful for preventing, ameliorating or treating medical
conditions, e.g. by protein or gene therapy. The genes are isolated from
a range of human tissues disclosed in the specification. The nucleic
acids, proteins, antibodies and (ant)agonists are useful in the
diagnosis, treatment and prevention of: (a) cancer, e.g. breast and
ovarian cancer, and other cancers of the adrenal gland, bone, bone
marrow, breast, gastrointestinal tract, liver, lung, or urogenital; (b)
immune disorders e.g. Addison's disease, allergies, autoimmune haemolytic
anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,
multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c)
cardiovascular disorders such as myocardial ischaemias; (d) wound healing
; (e) neurological diseases e.g. cerebral anoxia and epilepsy; and (f)
infectious diseases such as viral, bacterial, fungal and parasitic
infections

Q   Sequence 712 BP; 241 A; 121 C; 149 G; 201 T; 0 U; 0 Other;

Query Match   65.7%; Score 549; DB 3; Length 712;
Best Local Similarity 87.8%; Pred. No. 1.6e-128;
Matches 649; Conservative 0; Mismatches 0; Indels 90; Gaps 1;

```
y    97  TTCGCCTCACCCTCCCAGTGCACTGAAGAAGTAACCGGGTCCAGACCCACGCGGCGGCGCC  156
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b    29  TTCGCCTCACCCTCCCAGTGCACTGAAGAAGTAACCGGGTCCAGACCCACGCGGCGGCGCC   88

y   157  AGTTCTCCCGGCGGGAAGGAAGAAAACCGCGCAGAGAGGCAGCAATGAATGTGATCACGAGGT  216
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b    89  AGTTCTCCCGGCGGGAAGGAAGAAAACCGCGCAGAGAGGCAGCAATGAATGTGATCACGAGGT  148

y   217  TAACCTCTTAGTGGAGGAAATTCATCGTTTGGGTTCAAAAAATGCTGATGGAAAGTTAAG  276
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b   149  TAACCTCTTAGTGGAGGAAATTCATCGTTTGGGTTCAAAAAATGCTGATGGAAAGTTAAG  208

y   277  CGTGAAATTTGGGGTCCTCTTCCGTGATGATAAATGCCAACCTCTTTGAAGCATTGGT  336
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
b   209  CGTGAAATTTGGGGTCCTCTTCCGTGATGATAAATGTGCCAACCTCTTTGAAGCATTGGT  268
```

The invention relates to polynucleotides encoding colon tumour proteins.
The polynucleotides and encoded polypeptides are useful in pharmaceutical compositions, such as vaccines, for the diagnosis, prevention, and treatment of colon cancer. Polynucleotide sequences may be used as hybridisation probes or primers, and in the design and preparation of ribozyme molecules for inhibiting expression of tumour polypeptides and proteins in tumour cells. The compositions are useful for stimulating an immune response against cancer, particularly for the immunotherapy of colon cancer, and as markers for the progression of cancer. ABK44450-ABK46237 represent coding sequences of human colon tumour proteins of the invention. Note: With the exception of SEQ ID No 1 and 2, the sequence data for this patent did not form part of the printed specification but was supplied by the European Patent Office

Sequence 552 BP; 190 A; 104 C; 75 G; 183 T; 0 U; 0 Other;

Query Match   65.0%;   Score 542.8; DB 6; Length 552;
Best Local Similarity   99.6%;   Pred. No. 5.2e-127;
Matches 544;   Conservative   0;   Mismatches   2;   Indels   0;   Gaps   0;

```
290 GTCCCTCCGTGATGATAAATGTGCCAACCTCTTTGAAGCAATGGTAGGAACTCTTAAA 349
    |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
552 GTCCCTCCGTGATGATAAATGTGCCAACCTCTTTGAAGCAATGGTAGGAACTCTTAAA 493

350 GCTGCAAAACGAAGGAAGATTGTAACATATCCAGGAGGAGCTGCTTCTCGAAGGTGTTCAT 409
    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
492 GCTGCAAAACGAAGGAAGATTGTAACATATCCAGGAGGAGCTGCTTCTCGAAGGTGTTCAT 433

410 GATGATGTTGACATTATATTACTGCAAGATTAAATGGTTTACATATCTTTATGTACTGC 469
    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
432 GATGATGTTGACATTATATTACTGCAAGATTAAATGGTTTACATATCTTTATGTACTGC 373

470 CATTTTTGTTTCTGTAAACTGGAATATAAAAGTGAAAGAACAAACATTTGAACATACTT 529
    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
372 CATTTTTGTTTCTGTAAACTGGAATATAAAAGTGAAAGAACAAACATTTGAACATACTT 313

530 AATGTGATTTTTATAGAACTTGTAAACGAAAGGAGGATTCATGTTTTAGAAGTTGTGTCCTT 589
    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
312 AATGTGATTTTTATAGAACTTGTAAACGAAAGGAGGATTCATGTTTTAGAAGCCTGTCCTT 253

590 TTTTATATCTTGAAAGAAAATCTATGTATGATGCCTATAAAAATAAATCCTATTATTTTTCT 649
    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
252 TTTTATATCTTGAAAGAAAATCTATGTATGATGCCTATAAAAATAAATCCTATTATTTTTCT 193

650 CAGGAATCTGGTTAGGAATTGCAGGCAATGAGATTTTTGCGGGGCAGGGATGGGAATGT 709
    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
192 CAGGAATCTGGTTAGGAATTGCAGGCAATGAGATTTTTGCGGGGCAAGGATGGGAATGT 133

710 TTGTTCATAAATAATTAGACATTTTCTATAGATATTTGACATTCTGCGAAAGCAACAAGC 769
    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
132 TTGTTCATAAATAATTAGACATTTTCTATAGATATTTGACATTCTGCGAAAGCAACAAGC 73

770 AAACTGAAGACCAACTCCTATGAGAAATATTATGATGTTTATGTAATAAAGACATGTAAC 829
    |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
72 AAACTGAAGACCAACTCCTATGAGAAATATTATGATGTTTATGTAATAAAGACATGTAAC 13

830 TGTCTT 835
    ||||||
12 TGTCTT 7
```

SUIT 6
Q57055
ABQ57055 standard; cDNA; 594 BP.
ABQ57055;
02-AUG-2002 (first entry)

Human colon cancer related nucleotide sequence SEQ ID NO:750.

Human; colon cancer; cancer; tissue profiling; forensic; mapping;

Query Match   53.9%;   Score 450.2; DB 6; Length 594;
Best Local Similarity   86.8%;   Pred. No. 1.3e-103;
Matches 488;   Conservative   0;   Mismatches   74;   Indels   0;   Gaps   0;

```
Qy   6 GAGGCTTGAGCGCCAGAAACACTTACTTTTCCCCTACCCTGCTCCTCCTCCCCAGGCC 65
        |||| |||||||| ||||||||||||||||||||||||||||||||||||||| ||||
Db   8 GGGTCTTGAGCGCGCAGAAACACTTACTTTTCCCCTACCCTGCTCCTCCTCCCCACAGCC 67

Qy  66 GTCTTTCTCTTTGCCTCAGCCACTTCCTTCTTCGCCTCACCCTCACCCTCCCCAGTGCACTGAAG 125
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  68 GTCTTTCTCTTTGCCTCAGCCACTTCCTTCTTCGCCTCACCCTCACCCTCCCCAGTGCACTGAAG 127

Qy 126 AAGGTAACCGGGTCCAGACCCAGCGCGGCGGCCGGCCCAGTTCTCCGGCGGCGGGAAAACCGCGCCA 185
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 128 AAGGTAACCGGGTCCAGACCCAGCGCGGCGGCCGGCCCAGTTCTCCGGCGGCGGGAAAACCGCGCCA 187

Qy 186 GAGAGGCAGCAATGAATGTGGATCACGAGGTTAACCTTCATTAGTTGGAGGAAATTCATCGTT 245
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 188 GAGAGGCAGCAATGAATGTGGATCACGAGGTTAACCTTCATTAGTTGGAGGAAATTCATCGTT 247

Qy 246 TGGGTTCAAAAAATGTTGATGGAAAGTTAAGCGTGAAATTTGGGGTCCTTCTTCCGTGATG 305
        |||||||||||||||| |||||||||||||||||||||||||||||||||||||||||||
Db 248 TGGGTTCAAAAAATGCTGATGGAAAGTTAAGCGTGAAATTTGGGGTCCTTCTTCCGTGATG 307

Qy 306 ATAAATGTGCCAACCTCTTTGAAGCATTGGTAGGAACTCTTAAAGCTGCAAAACGAAGGA 365
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 308 ATAAATGTGCCAACCTCTTTGAAGCATTGGTAGGAACTCTTAAAGCTGCAAAACGAAGGA 367

Qy 366 AGATTGTAACATATCCAGGAGGAGCTGCTTCTCGAAGGTGTTCTGCATGATGATGTTGACATTA 425
```

```
Matches   453;   Conservative   0;   Mismatches   8;   Indels   2;   Gaps   1;

Qy      6 GAGGCTTGAGCGCAGAAACACTTACTTTTCCCCCTGCTCCTCCTCCCTCCTCCACAGCC  65
          ||||||||||||||||||||||||||||||||||||||||||||  ||||||||||||||
Db     11 GGGTCTTGAGCGCAGAAACACTTACTTTTCCCCCTACCTTCCTCCCTCCTCCCCACAGCC  70

Qy     66 GTCTTTCTCTTGCCTCAGCCACTTCCTTCCTTCCTGCCTCACCCTCCCAGTGCACTGAAG  125
          ||||||||||||||||||||||||| |||||||||||||||||||||||||||||||||||
Db     71 GTCTTTCTCTTGCCTCAGCCACTTCCTTCCTTCCTGCCTCACCCTCCCAGTGCACTGAAG  130

Qy    126 AAGGTAACCGGGTCCAGACCCACCGCGGCCAGTTCTCCGGCGGGAAGAAAACCGCGCA  185
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    131 AAGGTAACCGGGTCCAGACCCACCGCGGCCAGTTCTCCGGCGGGAAGAAAACCGCGCA  190

Qy    186 GAGAGGCAGCAATGAATGTGGATCACGAGGTTAACCTCTTAGTGCAGGAAATTCATCGTT  245
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    191 GAGAGGCAGCAATGAATGTGGATCACGAGGTTAACCTCTTAGTGCAGGAAATTCATCGTT  250

Qy    246 TGGGTTCAAAAAATGCTGATGGAAAGTTAAGCGTGAAATTTGGGGTCCTTCCGTGATG  305
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    251 TGGGTTCAAAAAATGCTGATGGAAAGTTAAGCGTGAAATTTGGGGTCCTTCCGTGATG  310

Qy    306 ATAAATGTGCCAACCTCTTTGAAGCATTGGTAGGAACTCTTAAAGCTGCAAAAACGAAGGA  365
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    311 ATAAATGTGCCAACCTCTTTGAAGCATTGGTAGGAACTCTTAAAGCTGCAAAAACGAAGGA  370

Qy    366 AGATTGTAACATATCCAGGAGAGCTGCTTCTGCAAGGTGTTCATGATGATGTTGACATTA  425
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    371 AGATTGTAACATATCCAGGAGAGCTGCTTCTACAAGGTGTTCATGATGATGTTGACATTA  430

Qy    426 TATTACTGCAA--GATTAATGTGGTTTACATATCTTTATGTAC  466
          |||||||||||   ||||||||||||||||||||||||||||||
Db    431 TATTACTGCAAAGAATAAATGNGGGTTACATATCTTTATGTAC  473
```

RESULT 8
```
ABN94189/c
ID   ABN94189 standard; DNA; 406 BP.
XX
AC   ABN94189;
XX
DT   13-AUG-2002  (first entry)
XX
DE   Gene #687 used to diagnose liver cancer.
XX
KW   Gene; liver cancer; ds; hepatocellular carcinoma; hepatotropic;
KW   metastatic liver tumour; cytostatic; expression profile; disease state;
KW   disease progression; drug toxicity; drug efficacy; drug metabolism.
XX
OS   Homo sapiens.
XX
PN   WO200229103-A2.
XX
PD   11-APR-2002.
XX
PF   02-OCT-2001; 2001WO-US030589.
XX
PR   02-OCT-2000; 2000US-0237054P.
XX
PA   (GENE-) GENE LOGIC INC.
XX
PI   Horne D,  Alvares C,  Peres-Da-Silva S,  Vockley JG;
XX
DR   WPI; 2002-426119/45.
XX
PT   Diagnosing and detecting the progression of liver cancer, hepatocellular
PT   carcinoma or metastatic liver tumor in a patient, involves detecting the
PT   level of expression of two or more genes in a liver tissue sample.
XX
PS   Claim 1; SEQ ID NO 687; 298pp; English.
XX
CC   The invention relates to a novel method for diagnosing and detecting the
CC   progression of liver cancer, hepatocellular carcinoma or metastatic liver
```

```
Qy    368 AGATTGTAACATATCCAGGAGAGCTGCTTCTACAAGGNGTTCATGATGATGTTGACATTA  427
          ||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
    426 TATTACTGCAAGATTAATGTGGTTTACATATCTTTATGTACTGTTTTGTTCTTGG  485
          ||||||||||||||||||||||||||||||||||||||||||||||||| |||||
    428 TATTACTGCAAGAATAATGNGGGTTACATATCTTTATGTACCCCGGCTCAATTGGATNCC  487

Qy    486 TAAACTGGAATATAAAGTGAAAGAACAAACATTGAACATACTTAATGTATTTTTATAGA  545
          |||||||||||||||||||||||||||||||||||||||||||||||||| ||||
    488 AATACTTGNCCACGGAACAAGGTANCCTAGGGATAACAACGCAAATNCTNTNTTTNAAAG  547

    546 ACTTTGTAAACGAAAGGAGAGATT  567
          ||||||||||||||||||| |||
    548 TCCATATTAACAATANGGGGTTT  569
```

RESULT 7
```
ABQ57053
ID   ABQ57053 standard; cDNA; 717 BP.
XX
AC   ABQ57053;
XX
DT   02-AUG-2002  (first entry)
XX
DE   Human colon cancer related nucleotide sequence SEQ ID NO:748.
XX
KW   Human; colon cancer; cancer; tissue profiling; forensic; mapping;
KW   genetic analysis; diagnostic; antisense therapy; gene; ss.
XX
OS   Homo sapiens.
XX
PN   WO200229086-A2.
XX
PD   11-APR-2002.
XX
PF   02-OCT-2001; 2001WO-US030732.
XX
PR   02-OCT-2000; 2000US-0237271P.
XX
PA   (FARB ) BAYER CORP.
XX
PI   Burgess C, Astle JH, Carroll E, Catino TJ, Dwivedi P, Molino GA;
PI   Thiaglingam A, Lewis ME;
XX
DR   WPI; 2002-426115/45.
XX
PT   New isolated nucleic acid that is differentially expressed in cancer
PT   tissues useful for determining the presence of colon cancer in a cell or
PT   tissue type, and in antisense therapy.
XX
PS   Claim 1; Fig 1; 796pp; English.
XX
CC   ABQ56306 to ABQ60787 represent isolated nucleic acids (I) differentially
CC   expressed in cancer tissues. ABB78893 to ABB79004 represent proteins
CC   encoded by the ABQ60776 to ABQ60787 nucleic acid sequences. (I) can be
CC   used in antisense therapy. An antibody immunoreactive with a polypeptide
CC   encoded by (I) is useful for detecting cancer in a patient sample, and
CC   for detecting the presence or absence of a polynucleotide encoded by a
CC   nucleic acid which hybridises to (I) in a cell. A probe/primer derived
CC   from (I) can be used for determining the presence of a nucleic acid which
CC   hybridises to (I), and for determining the phenotype of cells in a sample
CC   of cells from a patient. (I) is useful for determining the presence of
CC   colon cancer in a cell or tissue type, for determining the presence or
CC   state of other type of cancer, in antisense therapy, to generate
CC   macroarrays on a solid surface, to identify a chromosome on which the
CC   corresponding gene resides, and in tissue profiling, forensics, genetic
CC   analysis, mapping and diagnostic applications. (I) can be used to raise
CC   antibodies, and to screen for peptide analogues and antagonists
XX
SQ   Sequence 717 BP; 192 A; 171 C; 155 G; 180 T; 0 U; 19 Other;
```

```
Query Match         52.3%;  Score 436.8; DB 6; Length 717;
Best Local Similarity   97.8%; Pred. No. 3.4e-100;
```

PA   (CORI-) CORIXA CORP.
XX
PI   Jiang Y, Harlocker SL, Secrist H;
XX
DR   WPI; 2002-114514/15.
XX
PT   Novel isolated colon tumor polynucleotide differentially expressed in
PT   colon tumor or colon metastatic tumor and polypeptides encoded by them,
PT   useful for inhibiting development of cancer in patient.
PS   Claim 1; SEQ ID NO 2101; 105pp; English.
XX
CC   ABL36412 to ABL38645 represent human colon tumour antigen cDNA clones (I)
CC   which were isolated from human colon tumour and colon metastatic tumour
CC   cDNA libraries. (I) have cytostatic activity and can be used in vaccine
CC   production. (I) can be used for stimulating and/or expanding T cells
CC   specific for a tumour protein on contact with the T cells. They are also
CC   useful for inhibiting the development of cancer in a patient. (I) can be
CC   used as probes or primers for nucleic acid hybridisation, for preparing
CC   mutant species primers, or primers for use in genetic constructions. (I)
CC   can be used in the diagnosis of a colon tumour
XX
SQ   Sequence 368 BP; 128 A; 65 C; 44 G; 128 T; 0 U; 3 Other;

Query Match   43.7%;   Score 365.2; DB 6; Length 368;
Best Local Similarity 99.2%; Pred. No. 3.2e-82;
Matches 365; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy   462 TGTACTGCCATTTTTGTTCTGGTAAACTGGAATATAAGTGAAAGAACAAACATTTGA 521
              || |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   368 TNHACTGCCATTTTTGTTCTGGTAAACTGGAATATAAGTGAAAGAACAAACATTTGA 309

Qy   522 ACATACTTAAATGTATTTTATAGAACTTGTAAACGAAAGGAGATTCATGTTTTAGAAGT 581
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   308 ACATACTTAAATGTATTTTATAGAACTTGTAAACNAAAGGAGATTCATGTTTTAGAAGT 249

Qy   582 CTGTCCTTTTTATATCTTGAAAGAAATCTATGTATGATGCTATAAAATAAATCCTATT 641
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   248 CTGTCCTTTTTATATCTTGAAAGAAATCTATGTATGATGCTATAAAATAAATCCTATT 189

Qy   642 ATTTTTCTCAGGAATCTGGTTAGGAATTGCAGGCAATGAGATTTTTTGCGGGGCAGGGAT 701
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   188 ATTTTTCTCAGGAATCTGGTTAGGAATTGCAGGCAATGAGATTTTTTGCGGGGCAGGGAT 129

Qy   702 GGGAATGTTGTTCATAATAATTAGACATTTTCTATAGATATTATATGATGTTTATGGAAAG 761
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   128 GGGAATGTTGTTCATAATAATTAGACATTTTCTATAGATATTATATGATGTTTATGGAAAG 69

Qy   762 CAACAAGCAAACTGAGACAACCAACTCCTATGAGAAATATTATGATGTTTATGAATAAAGA 821
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   68 CAACAAGCAAACTGAGACAACCAACTCCTATGAGAAATATTATGATGTTTATGTAATAAAGA 9

Qy   822 CATGTAAC 829
              |||||||
Db   8   CATGTAAC 1

RESULT 10
ABK11085
ID   ABK11085 standard; cDNA; 780 BP.
XX
AC   ABK11085;
XX
DT   05-JUN-2002 (first entry)
XX
DE   Rat Progression Suppressed Gene 13 (rPSGen 13), cDNA.
XX
KW   Rat; Progression Suppressed Gene 13; rPSGen 13; cytostatic; cancer;
KW   blood vessel; nasopharyngeal tumour; thyroid tumour; leukaemia; lymphoma;
KW   breast; lung; prostate; ovary; colon; gene; ss.
XX
OS   Rattus sp.
XX

tumour in a patient, and differentiating metastatic liver cancer from
hepatocellular carcinoma in a patient, involving detecting the level of
expression of two or more genes represented in ABN93503-ABN97455 in a
tissue sample. The method of the invention has hepatotropic, and
cytostatic activity. The method is useful for diagnosing and detecting
the progression of liver cancer, hepatocellular carcinoma and metastatic
liver carcinoma in a patient. The method is useful for identifying
expression profiles which serve as useful diagnostic markers as well as
markers that can be used to monitor disease states, disease progression,
drug toxicity, drug efficacy and drug metabolism. Note: The sequence data
for this patent did not form part of the printed specification, but was
obtained in electronic format directly from WIPO at
ftp.wipo.int/pub/published_pct_sequences

Sequence 406 BP; 140 A; 68 C; 48 G; 150 T; 0 U; 0 Other;

Query Match   45.8%;   Score 382.4; DB 6; Length 406;
Best Local Similarity 99.7%; Pred. No. 1.5e-86;
Matches 383; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

452 CATATCTTATGTACTGCCATTTTTGTTCTGGTAAACTGGAATATAAAGTGAAAGAAC 511
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
406 CATATCCTTATGTACTGCCATTTTTGTTCTGGTAAACTGGAATATAAAGTGAAAGAAC 347

512 AAACATTTGAACATACTACTTAAATGTATTTTATAGAACTTGTAAACGAAAGGAGATTCATG 571
      |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
346 AAACATTTGAACATACTACTTAAATGTATTTTATAGAACTTGTAAACGAAAGGAGATTCATG 287

572 TTTTAGAAGTCTGTCCTTTTTATATCTTGAAAGAAAATCTANGTATGATGCTATAAAAT 631
      ||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
286 TTTTAGAAGTCTGTCCTTTTTATATCTTGAAAGAAAATCTATGTATGATGCTATAAAAT 227

632 AAATCCTATTATTTTTCTCAGGAATCTGGTTAGGAATTGCAGGCAATGAGATTTTTTGCG 691
      |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
226 AAATCCTATTATTTTTCTCAGGAATCTGGTTAGGAATTGCAGGCAATGAGATTTTTTGCG 167

692 GGGCAGGGATGGGAATGTTGTTCATAAATAATAGACATTTTCTATAGATATTTGACAT 751
      |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
166 GGGCAGGGATGGGAATGTTGTTCATAAATAATAATTAGACATTTTCTATAGATATTTGACAT 107

752 TCTGCGGAAAGCAACAAGCAAACTGAAGACAACCAACTCCTATGAGAAATATTATGATGTTTAT 811
      |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
106 TCTGCGGAAAGCAACAAGCAAACTGAAGACAACCAACTCCTATGAGAAATATTATGATGTTTAT 47

812 GTAATAAAGACATGTAACTGTCTT 835
      ||||||||||||||||||||||||
46 GTAATAAAGACATGTAACTGTCTT 23

SULT 9
L38512/c
ABL38512 standard; cDNA; 368 BP.

ABL38512;

08-APR-2002 (first entry)

Human colon tumour antigen polynucleotide SEQ ID NO:2101.

Homo sapiens.

Human; colon cancer; colon tumour antigen; cytostatic; vaccine;
colon tumour metastatic antigen; diagnosis; gene; ss.

WO200196388-A2.

20-DEC-2001.

08-JUN-2001; 2001WO-US018557.

09-JUN-2000; 2000US-0210899P.
20-FEB-2001; 2001US-0270216P.

```
Key       Location/Qualifiers
CDS       170..415
          /*tag= a
          /product= "Progression suppressed gene 13 protein"

WO20021 6419-A2.

28-FEB-2002.

27-AUG-2001; 2001WO-US026795.

25-AUG-2000; 2000US-00648310.

(UYCO ) UNIV COLUMBIA NEW YORK.

Fisher PB, Kang D, Su Z;

WPI; 2002-280914/32.
P-PSDB; AAU76532.

New rat and human Progression Suppressed Gene 13 for preventing the
growth of cancer cells and/or new blood vessels, and for treating
patients suffering from a cancer.

Claim 2; Fig 1; 53pp; English.

The invention relates to novel isolated nucleic acids which encode a rat
or human Progression Suppressed Gene 13 (PSGen 13) protein. The nucleic
acids are useful for preventing the growth of cancer cells and/or new
blood vessels, and for treating patients suffering from a cancer, e.g.
nasopharyngeal tumour, thyroid tumour, leukaemia, lymphoma, or cancer of
the breast, lung, prostate, ovary or colon. PSGen 13 may also be used to
suppress the transformed phenotype of a malignant cell. Administration of
PSGen 13 gene or protein may result in a decrease in tumour mass, number
of cancer cells, serum tumour marker, tumour metastasis, vascularisation,
perfusion, or rate of tumour growth, improved clinical symptoms, and/or
increased patient survival. The present sequence represents the coding
sequence of rat Progression Suppressed Gene 13 (rPSGen 13)
```

```
Sequence 780 BP; 223 A; 151 C; 187 G; 219 T; 0 U; 0 Other;

Query Match     40.1%;   Score 335.2; DB 6; Length 780;
Best Local Similarity   69.1%;   Pred. No. 1.6e-74;
Matches 539; Conservative   0; Mismatches 218; Indels 23; Gaps 5;

   43 CCTGCTCCTCCTCCTCCACAGCCGTCTTTCTCTTTGCCTCAGCCACTTCCTTCCTTCCGCC 102
      ||||| |||| |||||  ||||||  ||||||| || || |||||||||||||||||
   16 CGTCCCCTCCCTTCCTCCACTGCAGCCTTTCTCTTCTTTCTTAGCCCTAGCCACTTCCTTCCTTCCTTCTTCTGCT 75

  103 TCACCCTCCCCAGTGCACTGAAGAAGGTAACCGGGTCCAGACCCACCGCGGGCCAGTTCT 162
      ||| |||| |||||| ||||| |||| |||||||| || ||||| ||||| |||||
   76 TGTTCCTCCCTAGGGCCGGGAAGTCGAGTGCAGGGTTCAGCAGTGTAGCAGGCAGCTC 135

  163 CCGGCGGGAAGCAAAACCGCGCAGAGAGGCAGCAGCAATGAATGATGATCACGAGGTTAACCT 222
      ||||| ||||||||| ||||| ||| |||| ||| ||||| ||||||
  136 TTCAGTGAAGAAGCAATCGGAGGGTCAGGGGTCAGCAATGAACGTGGAGCATGAGGTTAACCT 195

  223 CTTAGTGGAGGAAATTCATCGTTGGGTTCAAAAAAATGCTGAGATGGAAAGTTAAGCGTGAA 282
      |||||||||||||| |||||||| ||||||| |||||||| ||||| |||||||||
  196 CCTGGTGGAGGAAATTCATCGTCTCGGGTTCCAAAAAATGCCGATGGGAAACTGAGTGTGAA 255

  283 ATTTGGGGTCCTCTTCCGTGATGATAAAATGTGCCAACCTCTTTGAAGCATTGTAGGAAC 342
      |||||||| ||||||||| ||||  ||||||| |||||||||||||| |||||
  256 GTTTGGGGTCCTCTCCAAGACGACAGAGATGTGCCAATCTCTTTGAAGCCGTTGGTGGGAAC 315

  343 TCTTAAAGCTGCAAAACGAAGGAGGAGATTGTAACATATCCAGGAGGAGAGCTGCTTCTGCAGG 402
      |||||||||||||| |||||||||||||||| ||||||||||||||||||||||||
  316 TCTGAAAGCCCGCAAAACGAAGGAAGGAGAGATTGTTACGTACGGAGGAGAGCTGCTTTTGCAAGG 375

  403 TGTTCATGATGATGTTGACATTATATTACTGCAAGATTAATTACTGGTTTACATATCTTTAT 462
      |||||||||||| |||||| ||||||  ||||||| |||| ||||| ||||| 
  376 TGTTCATGATGATGTTGACATTGTATTGCTGCAAGATTAATGTGGTTTTGCAGATCTGGGG 435
```

```
Qy  GTACTGCCATTTTTGTTTCTGGTAAACTGGAAT-ATAAAGTGAAAGAACAAACATTTGA 521
    |||                         ||||||||||||||||||||
Db  GTA---------------TCTGGTAAACTGGAATAATTAAGTTAAAGACAAACAT---G 477

Qy  ACATACTAATGTATTTTTATAGAACTTGTAAAGAAGAAGAGAGATTCATGTTTTAGAAGT 581
    |||||||||||||||||||||||||||   || ||||||  ||||||
Db  AAGTTCCTATGTATTTTTATAGACCTTGTAAACAAAAGGGGA--CTTGTTGAGAAGTC 535

Qy  CGTGTCCTTTTTATATCTGAAAGAAAATCATGTAAGATGGTCTATAAAATAAATCCTATT 641
    |||||||                  ||||||||||| |||| |||||||||
Db  CTGTTTTTATACCTTGGAGCAAAACATTACAATGTAAAATAAACAAAACCTGTTATTTT 595

Qy  ATTTTTCTCAGGAATCTGGTTAGGAATTGCAGGCAATGAGAGATTTTTGCGGCGCAGGGAT 701
    ||||||      |||   |||  ||  |||||||   || 
Db  TTTTTTCTTAAGAGGTAATCGGGAGACGTAGGCAATAAAAATGTTTTCAGAGGTGCGAAA 655

Qy  GGGAATGTTTGTTCATAAATAATATAGACATTTTCTATAGATATTTGACATTCTGCGAAAG 761
    ||||| ||||                |  ||||  |||||||||
Db  AAGCTTTTGTTTTCTTAAACCATTCCTTAGTCT--CTGCCACACTTGACACTCCGTCAAAG 713

Qy  CAACAAGCAAACTGAAGACCAACTCCTATGAGAAATATTATGATGTTTATGTTAATAAAGA 821
    |||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db  TGAGAAGCGAACTAAAGACCAACTGCGGTGGAAAATATTATGTTTATGTTAATAAAAAAAA 773
```

```
Key       Location/Qualifiers
CDS       170..415
          /*tag= a
          /product= "Progression suppressed gene 13 protein"

RESULT 11
AAZ21517
ID    AAZ21517 standard; cDNA; 800 BP.
XX
AC    AAZ21517;
XX
DT    01-DEC-1999  (first entry)
XX
DE    Progression suppressed gene 13 (PSGen13).
XX
KW    Progression suppressed gene; PSGen; progression elevated gene; PEGen;
KW    tumour; reciprocal subtraction differential RNA display; RSDD;
KW    differential expression; gene cloning; cancer; ss.
XX
OS    Rattus sp.
XX
FH    Key          Location/Qualifiers
FT    CDS          170..448
FT                 /*tag= a
FT                 /product= "PSGen13"
FT                 /note= "Progression suppressed gene 13 protein"
XX
PN    WO9943844-A1.
XX
PD    02-SEP-1999.
XX
PF    26-FEB-1999;  99WO-US004323.
XX
PR    27-FEB-1998;  98US-00032684.
PR    03-NOV-1998;  98US-00185115.
PR    23-NOV-1998;  98US-00197889.
XX
PA    (UYCO ) UNIV COLUMBIA NEW YORK.
XX
PI    Fisher PB;
XX
DR    WPI; 1999-550872/46.
DR    P-PSDB; AAY39325.
XX
PT    Identifying nucleic acids differentially expressed between two samples,
PT    particularly sequences involved in tumor progression.
XX
PS    Claim 21; Fig 35B; 110pp; English.
XX
CC    This sequence is the progression suppressed gene 13 (PSGen13). This gene
CC    has suppressed expression in progressed tumour cells. PSGen13 was
CC    identified using new methods for identifying nucleic acids differentially
CC    expressed between two samples. The method involves performing reciprocal
```

subtraction differential RNA display (RSDD) between the two samples to
generate two subtraction samples are amplified
and compared to identify those nucleic acids that are differentially
expressed. The method is used to identify and clone differentially
expressed genes, particularly those with increased or reduced expression
during tumour cell progression, e.g. progression suppressed genes (PSGen)
and progression elevated genes (PEGen). The method reduces the complexity
of the band pattern produced in conventional differential RNA display
(where bands may be obscured, resulting in false positive signals) since
most bands common to both samples are eliminated, allowing identification
and cloning of genes displaying anticipated differential expression. RSDD
requires only a single anchored primer for amplification and reamplified
cDNA can be analysed by reverse Northern blotting

Sequence 800 BP; 243 A; 153 C; 185 G; 219 T; 0 U; 0 Other;

Query Match   37.1%; Score 310; DB 2; Length 800;
Best Local Similarity 68.7%; Pred. No. 3.8e-68;
Matches 537; Conservative   0; Mismatches 220; Indels 25; Gaps   7;

```
 43 CCTGCTCCTCCCTCCTCCACAGCCGTCTTCCTTTGCCTCCGCCACTTCCTTCCTTCCGCC 102
    |||||||| ||||| |||| |||  |||  ||||  |||||||  ||| |||||||||||
 16 CGTCCCCTCCCTTCTCCACTGCAGCCTTTCTTCTTAGCCCGAACCACTTCCTTCTTCTGCT 75

103 TCACCCTCCCCAGTGCACTGAAGAAGTAACCGGGTCCAGACCCACGCGGCGCCAGTTCT 162
    |||||||||||
 76 TGTTCCTCCCTAGGGCGCGGAAGCTCAGTGCAGGGTTCAGCACCCCACGCGCGGCGAGCTC 135

163 CCGGCGGAAGGAAAAACCGCGCAGAGAGGCAGCAATGAATGTGGATCACGAGGTTAACCT 222
    |||||||||||||
136 TTCAGTGAAGAAGGAAGCAATCGGAGGGTCAGCAATGAAGCTGAGCAATGAGAGGTTAACCT 195

223 CTTAGTGGAGGAAAATTCATCGTTTGGGTTCAAAAAATGCTGATGGAAAGTAAGCGTGAA 282
    |||||||||||||
196 CCTGGTCGAGGAAATTCATCGTCTGGGTTCCAAAAATGCCGATGGGAAACTGAGTGTGAA 255

283 ATTTGGGGTCCTCTTCCGTGATGATAAATGCCAACCTCTTTG-AAGCATTGGTAGGAA 341
    |||||||||||
256 GTTTGGGGTCCTTCTTCCAAGACGACAGATGTGCCAATCTCTTTGAAAACCGTTGGTGGGAA 315

342 CTCTTAAAG-CTGCAAAACGAAGGAGATTGTAACATATCCAGGAGAGCGCTTCTGCAA 400
    ||||||||
316 CTCTGAAAGCCCGCAAAAACGAAGGAGAGATTGTTACGTACGCAGAGAGCGTCTTTTGCA 375

401 GGTGTTCATGATGATCGTTGACATTATATACTGCAAGATAAATGTGGTTTACATATCTTT 460
    |||||||||||
376 GGTGTTCATGATGATGTTGACATTGTATTGTCGCAAGATTAATGTGGTTTGCAGATCTGG 435

461 ATGTACTGCCATTTTTGTTTCTGGTAAACGGGAAT-ATAAAGTGAAAGAACAAACATTT 519
    ||||||||||
436 GGGTA----------TCTGGTAAACTGGAATAATTAAAGGACAAACAT-- 478

520 GAACATACTTAATGATTTTTATAGAACTTGTAAACGAAAGGAGAGATTCATGTTTTAGAAG 579
    |||||||||||
479 -GAAGTTCCTTATGTATTTTTATAGACCTTTGTAAACAAAAGGGGA--CTTGTTGAGAAG 535

580 GTCTGTCCTTTTTATATCTTGAAAGAAAATCTATGTATGATGCTATAAAATAATCCTA 639
    |||||||||||
536 TCCTGTTTTTTATACCTTGGAGCAAAACATTACAATGTAAAAATAAACAAAACCTGTTATT 595

640 TTATTTTTCTCAGGAATCTGGTTAGGAATTGCAGGCAATGAGATTTTTGCGGGGCAGGG 699
    |||||||||||
596 TTTTTTTTCTTAAGAGGTAATCGGGAGACGTAGGCAATAAAAATGTTTTCCAGGGTGCCA 655

700 ATGGGAATGTTGTTCATAAATAATTAGAACATTTGTTATAGATATTTGACATTCTGCAA 759
    |||||||||||
656 AAAAGCTTTTGTTTTTTTAAACCATTCTTAGTGTCT--CTGCCACACTTGACACTCCGTCAA 713

760 AGCAACAAGCAAACTGAAGACCAACTCCTATGAGAAATATTATGATGTTTATGTAATAAA 819
    |||||||||||
714 AGTGAGAAGCGGAACTAAAGAACCAACTCGCGTGGAAAAATATTTATGTTTTATGTAATAAAAA 773

820 GA 821
```

RESULT 12
AAC30197
ID   AAC30197 standard; cDNA; 296 BP.
XX
AC   AAC30197;
XX
DT   06-OCT-2000 (first entry)
XX
DE   Human secreted protein 5' EST, SEQ ID NO: 34272.
XX
KW   Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
KW   gene therapy; chromosome mapping; ss.
XX
OS   Homo sapiens.
XX
PN   EP1033401-A2.
XX
PD   06-SEP-2000.
XX
PF   21-FEB-2000; 2000EP-00200610.
XX
PR   26-FEB-1999; 99US-0122487P.
XX
PA   (GEST) GENSET.
XX
PI   Dumas Milne Edwards J, Duclert A, Giordano J;
XX
DR   WPI; 2000-500381/45.
XX
PT   New nucleic acid that is a 5' expressed sequence tag (5' EST) for
PT   obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
PT   diagnostic, forensic, gene therapy and chromosome mapping procedures.
XX
PS   Claim 1; SEQ ID NO 34272; 71pp + Sequence Listing; English.
XX
CC   The present sequence is one of a large number of 5' ESTs derived from
CC   mRNAs encoding secreted proteins. No ORF has yet been conclusively
CC   identified within the present sequence. The 5' ESTs were prepared from
CC   total human RNAs or polyA+ RNAs derived from 30 different tissues. EST
CC   sequences usually correspond mainly to the 3' untranslated region (UTR)
CC   of the mRNA because they are often obtained from oligo-dT primed cDNA
CC   libraries. Such ESTs are not well suited for isolating cDNA sequences
CC   derived from the 5' ends of mRNAs and even in those cases where longer
CC   cDNA sequences have been obtained, the full 5' UTR is rarely included. 5'
CC   ESTs are derived from mRNAs with intact 5' ends and can therefore be used
CC   to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used in
CC   diagnostic, forensic, gene therapy and chromosome mapping procedures.
CC   They are used to obtain upstream regulatory sequences and to design
CC   expression and secretion vectors
XX
SQ   Sequence 296 BP; 105 A; 32 C; 57 G; 100 T; 0 U; 2 Other;

Query Match   33.5%; Score 279.6; DB 3; Length 296;
Best Local Similarity 99.3%; Pred. No. 1.2e-60;
Matches 290; Conservative 1; Mismatches   0; Indels   1; Gaps   1;

```
Qy  487 AAACTGGAATATAAAGTGAAAGAACAAACATTGAACATACTTAATGTATTTTATAGAA 546
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 AAACTGGAATATAAAGTGAAAGAACAAACATTGAACATACTTAATGTATTTTTATAGAA 60

Qy  547 CTTTGTAAACGAAAGGAGATTCATGTTTAGAAGTCTGTCCTTTTTATATCTTGAAAGA 606
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 CTTTGTAAACGAAAGGAGATTCATGTTTAGAAGTCTGTCCTTTTTTATATCTTGAAAGA 120

Qy  607 AAATCTATGTATGATGCTATAAAATAAATCCTATTATTTTTCTCAGGAATCTGGTTAGGA 666
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 AAATCTATGTATGATGCTATAAAATAAATCCTATTATTTTTCTCAGGAATCTGGTTAGGA 180

Qy  667 ATTGCAGGCAATGAGATTTTTGCGGGGCAGGGGATGGAATGTTTTGTTCATAATAATTA 726
```

```
Db  499 TGCCTCAGCCACTTCCTTCCTTCCTGCCTCACCCTCCCCAGTGCATTGAAGAAGGTAACCG 440
Qy  136 GGTCCAGACCCACGCGCGCGCGCCAGTTCTCCGGCGGGAACGAAAACCGCCGCCAGAGA----- 189
Db  439 GGTCCAGACCCACCGCGCGGCGGCCAGTTCTCCGGCGGCGGAAGGAAAAACCGCCGAGAGGCCCT 380
Qy  190 ------------------------------------------------------------- 189
Db  379 CATCTGCCTTAGAATGTGAGATCCCCAGGGCGAGATCCTGTGCTACACACCTTTTGTAT 320
Qy  190 ----------------------------------GGCAGCAATGAATGTGGAT 208
Db  319 TTTTGGTGCTGGCCAGGATACCATGAAGAGGACCTACGCATGGCAGCAATGAATGTGGAT 260
Qy  209 CACGAGGTTAACCTCTTAGTGGAGGAAATTCATCGTTGGGTTCAAAAAATGCTGATGGA 268
Db  259 CACGAGGTTAACCTCTTAGTGGAGGAAATTCATCGTTTGGGTTCAAAAAATGCTGATGGA 200
Qy  269 AAGTTAAGCGTGAAATTTGGGTCCTTCCGTGATGATAAATGTGCCAACCTCTTTGAA 328
Db  199 AAGTTAAGCGTGAAATTTGGGTCCTTCCTCCGTGATGATAAATGTGCCAACCTCTTTGAA 140
Qy  329 GCATTGGTAGGAACTCTTAAAGCTGCAAAACGAAGGAAGATTGTAACATATCCAGGAGAG 388
Db  139 GCATTGGTAGGAACTCTTAAAGCTGCAAAACGAAGGAAGATTGTAACATATCCAGGAGAG 80
Qy  389 CTGCTTCTGCAAGGTGTTCATGATGGTTCATGATGGTTGACATTATTATTACTGCAAGATTAATGTGGT 448
Db  79  CTGCTTCTGCAAGGTGTTCATGATGATGTTGACATTATTATTACTGCAAGATTAATGTGGT 20
Qy  449 TTACATATCTTTATGTAC 466
Db  19  TTACATATCTTTATGTAC 2
```

RESULT 14
ACH38863
ID ACH38863 standard; cDNA; 458 BP.
AC ACH38863;
DT 13-OCT-2003 (first entry)
DE Human foetal brain cDNA #230.
KW Human; ss; sequencing by hybridisation; SBH; expressed sequence tag; EST; genome mapping; biodiversity; genetic disorder.
OS Homo sapiens.
PN US2003073623-A1.
PD 17-APR-2003.
PF 30-JUL-2001; 2001US-00918995.
PR 30-JUL-2001; 2001US-00918995.
PA (DRMA/) DRMANAC R T.
PA (LABA/) LABAT I.
PA (STAC/) STACHE-CRAIN B.
PA (DICK/) DICKSON M C.
PA (JONE/) JONES L W.
PI Drmanac RT, Labat I, Stache-Crain B, Dickson MC, Jones LW;
DR WPI; 2003-615964/58.
PT New polynucleotide sequences obtained from various cDNA libraries, useful as hybridization probes, as oligomers for PCR, for chromosome and gene mapping, in the recombinant production of protein, or in generating antisense DNA or RNA.

(second column / earlier part)

```
181 ATTGCAGGCAATGACAATTTTTGCGGGGCAGGGATGGGAATGTTGTTCATTAAAATAATTA 240
727 GACATTTTCTATAGATATTGACATTCTGCGAAAGCAACAAGCAAAACTGAAG 778
241 GACATTTTTCTATAGATATTGACATTCTCGCGAAA-SAACAAGCAAAACTGAAG 291
```

RESULT 13
ABQ56862/c
ABQ56862 standard; cDNA; 627 BP.
ABQ56862;
02-AUG-2002 (first entry)
Human colon cancer related nucleotide sequence SEQ ID NO:557.
Human; colon cancer; cancer; tissue profiling; forensic; mapping; genetic analysis; diagnostic; antisense therapy; gene; ss.
Homo sapiens.
WO2002229086-A2.
11-APR-2002.
02-OCT-2001; 2001WO-US030732.
02-OCT-2000; 2000US-0237271P.
(FARB) BAYER CORP.
Burgess C, Astle JH, Carroll E, Catino TJ, Dwivedi P, Molino GA; Thiaglingam A, Lewis ME;
WPI; 2002-426115/45.
New isolated nucleic acid that is differentially expressed in cancer tissues useful for determining the presence of colon cancer in a cell or tissue type, and in antisense therapy.

Claim 1; Fig 1; 796pp; English.

ABQ56306 to ABQ60787 represent isolated nucleic acids (I) differentially expressed in cancer tissues. ABB78993 to ABB79004 represent proteins encoded by the ABQ60776 to ABQ60787 nucleic acid sequences. (I) can be used in antisense therapy. An antibody immunoreactive with a polypeptide encoded by (I) is useful for detecting cancer in a patient sample, and for detecting the presence or absence of a polynucleotide encoded by a nucleic acid which hybridises to (I) in a cell. A probe/primer derived from (I) can be used for determining the presence of a nucleic acid which hybridises to (I), and for determining the phenotype of cells in a sample of cells from a patient. (I) is useful for determining the presence of colon cancer in a cell or tissue type, for determining the presence or state of other type of cancer, in antisense therapy, to generate macroarrays on a solid surface, to identify a chromosome on which the corresponding gene resides, and in tissue profiling, forensics, genetic analysis, mapping and diagnostic applications. (I) can be used to raise antibodies, and to screen for peptide analogues and antagonists

Sequence 627 BP; 171 A; 151 C; 153 G; 147 T; 0 U; 5 Other;
Query Match 33.5%; Score 279.4; DB 6; Length 627;
Best Local Similarity 76.9%; Pred. No. 1.8e-60;
Matches 429; Conservative 0; Mismatches 19; Indels 110; Gaps 3;

```
Qy  19  AGAAAACAACTTACTTTTCCCCCTACCCTGCTCCTCCTCCTCCTCCACAGCCG--TCTTTCTCTT 76
Db  559 AGAAAACATTACTTTTTCCCCNTACCCTGGTCTTCTTCCTCCCCACAGCNGTTTTTTTTNTTT 500
Qy  77  TGCCTCAGCCAC-TTCCTTCCTTCCTGCCTCACCCTCCCCAGTGCACTGAAGAAGGTAACCG 135
```

Claim 1; SEQ ID NO 26075; 44pp; English.

The invention relates to an isolated polynucleotide comprising any one of
38043 cDNA sequences, appearing as ACH12789-ACH50831, whose sequence was
determined by the technique of SBH (sequencing by hybridisation). Also
included is a purified polypeptide comprising a sequence corresponding to
a reading frame of the novel polynucleotide. The nucleic acid sequences
are useful in diagnostics as expressed sequence tags (EST) for
identifying expressed genes or for physical mapping of the human genome,
in forensics, in assessing biodiversities, or in identifying mutations
responsible for genetic disorders and other traits. The nucleotide
sequences are also useful as hybridisation probes, as oligomers for PCR,
for chromosome and gene mapping, in the recombinant production of
protein, or in generating antisense DNA or RNA. The purified polypeptide
is useful for generating antibodies specific for it. The present sequence
is one of the 38043 isolated cDNA/EST sequences. Note: The sequence data
for this patent did not form part of the printed specification, but was
obtained in electronic format directly from USPTO at
seqdata.uspto.gov/sequence.html?DocID=20030073623

Sequence 458 BP; 112 A; 86 C; 149 G; 107 T; 0 U; 4 Other;

Query Match   28.4%;   Score 237;   DB 8; Length 458;
Best Local Similarity   100.0%;   Pred. No. 8.4e-50;
Matches 237;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

```
      188 GAGGCAGCAATGAATGTGGATCACGAGGTTAACCTCTTAGTGGAGGAAATTCATCGTTTG 247
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
      222 GAGGCAGCAATGAATGTGGATCACGAGGTTAACCTCTTAGTGGAGGAAATTCATCGTTTG 281

      248 GGTTCAAAAAATGCTGATGGAAAGTTAAGCGTGAAATTTGGGGTCCTCTTCCGTGATGAT 307
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
      282 GGTTCAAAAAATGCTGATGGAAAGTTAAGCGTGAAATTTGGGGTCCTCTTCCGTGATGAT 341

      308 AAATGTGCCAACCTCTTGAAGCATTGGTAGGAACTCTTAAAGCTGCAAAACGAAGGAAG 367
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
      342 AAATGTGCCAACCTCTTGAAGCATTGGTAGGAACTCTTAAAGCTGCAAAACGAAGGAAG 401

      368 ATTGTAACATATCCAGGAGAGCTGCTTCTGCAAGGTGTTCATGATGATGTTGACATT 424
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
      402 ATTGTAACATATCCAGGAGAGCTGCTTCTGCAAGGTGTTCATGATGATGTTGACATT 458
```

SULT 15
A11066
ABA11066 standard; cDNA; 351 BP.

ABA11066;

23-JAN-2002   (first entry)

Human nervous system related polynucleotide SEQ ID NO 73.

Human; nootropic; neuroprotective; cytostatic; dermatological; virucide;
immunosuppressive; antiinflammatory; anti-HIV; antibacterial; vulnerary;
antiparkinsonian; antisickling; antianaemic; antiarthritic; cancer;
antirheumatic; hepatotropic; cerebroprotective; antiinflammatory;
antiallergic; antidiabetic; antiulcer; anticonvulsant; antifungal;
antiparasitic; cardiant; immune disorder; cardiovascular disorder;
neurological disease; infection; nephrotropic; gene therapy; vaccine; ss.

Homo sapiens.

WO200159063-A2.

16-AUG-2001.

17-JAN-2001; 2001WO-US001334.

31-JAN-2000; 2000US-0179065P.
04-FEB-2000; 2000US-0180628P.
24-FEB-2000; 2000US-0184664P.
02-MAR-2000; 2000US-0186350P.

PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226868P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0204437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.

CC   The invention relates to novel genes (ABA11004-ABA21534) and proteins
CC   (ABB14678-ABB18001) useful for preventing, treating or ameliorating
CC   medical conditions e.g. by protein or gene therapy. The genes are
CC   isolated from a range of human tissues disclosed in the specification.
CC   The nucleic acids, proteins, antibodies and (ant)agonists are useful in
CC   the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and
CC   ovarian cancer and other cancers of the adrenal gland, bone, bone marrow,
CC   breast, gastrointestinal tract, liver, lung, or urogenital; (b) immune
CC   disorders e.g. Addison's disease, allergies, autoimmune haemolytic
CC   anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,
CC   multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c)
CC   cardiovascular disorders such as myocardial ischaemias; (d) wound healing
CC   ; (e) neurological diseases e.g. cerebral anoxia and epilepsy; and (f)
CC   infectious diseases such as viral, bacterial, fungal and parasitic
CC   infections. Note: The sequence data for this patent did not form part of
CC   the printed specification, but was obtained in electronic format directly
CC   from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 351 BP; 104 A; 72 C; 103 G; 72 T; 0 U; 0 Other;

Query Match   25.9%; Score 216.4; DB 5; Length 351;
Best Local Similarity   96.2%; Pred. No. 1.2e-44;
Matches 254; Conservative   0; Mismatches   6; Indels   4; Gaps   3;

Qy   114 AGTGCACTGAAGAAGTAACCGGGTCCAGACCCCAGCCGCGGCCACCCCAGTTCTCCGGCCGGGAAG 173
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   90 AGTGCACTGAAGAAGTAACCGGGTCCAGACCCCAGCCGCGGCCACCCCAGTTCTCCGGCCGGGAAG 149

Qy   174 GAAAACCGCGCAGAGAGGCAGCAATGAATGGATCACGAGGTTAACCTCTTAGTGGAGG 233
          |||||||||||||||||||||||||||||| |||||||||||||||||||||||||||
Db   150 GAAAACCGCGCAGAGAGGCAGCAATGAATGTGGATCACGAGGTTAACCTCTTAGTGGAGG 209

Qy   234 AAATTCATCGTTTGGGTTCAAAAAATGCTGATGGAAAGTTAAGCGTGAAATTTGGGGT-- 291
          |||||||||||||||||||||||||| |||||||||||||||||||||||||||||||
Db   210 AAATTCATCGTTTGGGTTCAAAAAATCTGATGGAAAGTTAAGCGTGAAATTTGGGGTTC 269

Qy   292 CCTCTTCCGTGATGATAAATGTGCCAACCTCTTGAAGCATTGGTAGGAACTCTTAAAGC 351
          | ||||||||||||||||||||||| ||||||||||||| ||||||||||||||||||
Db   270 CTCTTCCGTGATGATAAATGTGCC-ACCTCTTGAAGCA-TGGTAGGAACCCCTTAAAGC 327

Qy   352 TGCAAAACGAAGGAAGATTGTAAC 375
          ||||||||||||||||| | ||||
Db   328 TGCAAAACGAAGGAAGATGTTAAC 351

Search completed: March 17, 2004, 12:34:42
Job time : 445 secs

~   13-OCT-2000; 2000US-0239935P.
~   13-OCT-2000; 2000US-0239937P.
~   20-OCT-2000; 2000US-0240960P.
~   20-OCT-2000; 2000US-0241785P.
~   20-OCT-2000; 2000US-0241786P.
~   20-OCT-2000; 2000US-0241787P.
~   20-OCT-2000; 2000US-0241808P.
~   20-OCT-2000; 2000US-0241809P.
~   20-OCT-2000; 2000US-0241826P.
~   20-OCT-2000; 2000US-0242221P.
R   01-NOV-2000; 2000US-0244617P.
R   08-NOV-2000; 2000US-0246474P.
R   08-NOV-2000; 2000US-0246475P.
R   08-NOV-2000; 2000US-0246476P.
R   08-NOV-2000; 2000US-0246477P.
R   08-NOV-2000; 2000US-0246478P.
R   08-NOV-2000; 2000US-0246523P.
R   08-NOV-2000; 2000US-0246524P.
R   08-NOV-2000; 2000US-0246525P.
R   08-NOV-2000; 2000US-0246526P.
R   08-NOV-2000; 2000US-0246527P.
R   08-NOV-2000; 2000US-0246528P.
R   08-NOV-2000; 2000US-0246532P.
R   08-NOV-2000; 2000US-0246609P.
R   08-NOV-2000; 2000US-0246610P.
R   08-NOV-2000; 2000US-0246611P.
R   08-NOV-2000; 2000US-0246613P.
R   17-NOV-2000; 2000US-0249207P.
R   17-NOV-2000; 2000US-0249208P.
R   17-NOV-2000; 2000US-0249209P.
R   17-NOV-2000; 2000US-0249210P.
R   17-NOV-2000; 2000US-0249211P.
R   17-NOV-2000; 2000US-0249212P.
R   17-NOV-2000; 2000US-0249213P.
R   17-NOV-2000; 2000US-0249214P.
R   17-NOV-2000; 2000US-0249215P.
R   17-NOV-2000; 2000US-0249216P.
R   17-NOV-2000; 2000US-0249217P.
R   17-NOV-2000; 2000US-0249218P.
R   17-NOV-2000; 2000US-0249244P.
R   17-NOV-2000; 2000US-0249245P.
R   17-NOV-2000; 2000US-0249264P.
R   17-NOV-2000; 2000US-0249265P.
R   17-NOV-2000; 2000US-0249297P.
R   17-NOV-2000; 2000US-0249299P.
R   17-NOV-2000; 2000US-0249300P.
R   01-DEC-2000; 2000US-0250391P.
R   01-DEC-2000; 2000US-0251160P.
R   05-DEC-2000; 2000US-0251030P.
R   05-DEC-2000; 2000US-0251988P.
R   05-DEC-2000; 2000US-0256719P.
R   06-DEC-2000; 2000US-0251479P.
R   08-DEC-2000; 2000US-0251856P.
R   08-DEC-2000; 2000US-0251868P.
R   08-DEC-2000; 2000US-0251869P.
R   08-DEC-2000; 2000US-0251989P.
R   08-DEC-2000; 2000US-0251990P.
R   11-DEC-2000; 2000US-0254097P.
R   05-JAN-2001; 2001US-0259678P.

X   (HUMA-) HUMAN GENOME SCI INC.
X
T   Rosen CA, Barash SC, Ruben SM;
X
R   WPI; 2001-541565/60.
R   P-PSDB; ABB14740.
X
T   Nucleic acids encoding 3224 human nervous system antigen polypeptides,
T   useful for preventing, diagnosing and/or treating nervous system cancers
    and metastases.
X
X   Claim 1; SEQ ID NO 73; 1701pp + Sequence Listing; English.

```
            GenCore version 5.1.6
        Copyright (c) 1993 - 2004 Compugen Ltd.

nucleic - nucleic search, using sw model

on:     March 17, 2004, 12:23:02 ; Search time 2821 Seconds
                                    (without alignments)
                                    8839.036 Million cell updates/sec

le:     US-09-648-310-3
fect score: 835
uence:  1 ggcacgaggcttgagcgcag.........taaagacatgtaactgtcctt 835

ring table: IDENTITY_NUC
        Gapop 10.0 , Gapext 1.0

rched:  27513289 seqs, 14931090276 residues

al number of hits satisfying chosen parameters:    55026578

imum DB seq length: 0
imum DB seq length: 2000000000

t-processing: Minimum Match 0%
              Maximum Match 100%
              Listing first 45 summaries

abase :  EST:*
    1:  em_estba:*
    2:  em_esthum:*
    3:  em_estin:*
    4:  em_estmu:*
    5:  em_estov:*
    6:  em_estpl:*
    7:  em_estro:*
    8:  em_htc:*
    9:  gb_est1:*
   10:  gb_est2:*
   11:  gb_htc:*
   12:  gb_est3:*
   13:  gb_est4:*
   14:  gb_est5:*
   15:  em_estfun:*
   16:  em_estom:*
   17:  em_gss_hum:*
   18:  em_gss_inv:*
   19:  em_gss_pln:*
   20:  em_gss_vrt:*
   21:  em_gss_fun:*
   22:  em_gss_mam:*
   23:  em_gss_mus:*
   24:  em_gss_pro:*
   25:  em_gss_rod:*
   26:  em_gss_phg:*
   27:  em_gss_vrl:*
   28:  gb_gss1:*
   29:  gb_gss2:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

sult No.	Score	% Query Match	Length	DB	ID	Description
1	805.4	96.5	845	11	AF161398	AF161398 Homo sapi
2	762.6	91.3	876	11	AF116682	AF116682 Homo sapi
3	750.6	89.9	816	13	BU849360	BU849360 AGENCOURT
4	737	88.3	797	12	BI859694	BI859694 60338 6283

```
c  5  715.6  85.7  749  12  BQ015110   BQ015110 UI-H-ED1-
   6  695.2  83.3  741   9  AV717724   AV717724 AV717724
   7  694.6  83.2  752   9  AV716363   AV716363 AV716363
   8  683.8  81.9  714   9  AV713584   AV713584 AV713584
   9  673.8  80.7  906  14  CD521898   CD521898 AGENCOURT
  10  671.4  80.4  676  13  BX110370   BX110370 BX110370
  11  667.4  79.9  717  13  BU98760    BU98760  AGENCOURT
  12  663    79.4  755  12  BG530613   BG530613 602560008
  13  657.8  78.8  765  13  BU93438    BU93438  AGENCOURT
  14  657.6  78.8  747  12  BG502291   BG502291 602550569
  15  657    78.7  684  12  BG744405   BG744405 602723144
  16  655.4  78.5 1126  12  BM542613   BM542613 AGENCOURT
  17  655    78.4  694  12  BG500857   BG500857 602547312
  18  652    78.1  698  10  BE891760   BE891760 601434485
c 19  651.6  78.0  678   9  AW027019   AW027019 wt71b02.x
c 20  650    77.8  674   9  AW027067   AW027067 wt71f11.x
  21  649.2  77.7  791  12  BG528919   BG528919 602579350
  22  645.6  77.3  721   9  AV713252   AV713252 AV713252
c 23  645.6  77.3  664   9  AW027058   AW027058 wt71e11.x
  24  642    76.9  812  14  CB958631   CB958631 AGENCOURT
c 25  642    76.9  666  12  BG745050   BG745050 602723144
c 26  642    76.9  682  12  BQ020407   BQ020407 UI-H-DP0-
  27  641.6  76.8  763   9  AV756659   AV756659 AV756659
c 28  638.4  76.5  653   9  AW024795   AW024795 wu69a04.x
  29  636.8  76.3  648  13  BQ777108   BQ777108 i143e02.y
  30  635.6  76.1  884  12  BG164526   BG164526 602342067
  31  624.4  74.8  773  12  BG499248   BG499248 602545693
  32  622.4  74.5  815  13  BU199110   BU199110 DCBCKCC05
  33  620.6  74.3  634   9  AI133406   AI133406 HA2013 Hu
  34  620.2  74.3 1095  12  BG114255   BG114255 602285822
c 35  617.8  74.0  640  13  BQ778215   BQ778215 i143e02.x
  36  613.8  73.5  641  12  BG563686   BG563686 602584452
  37  608.4  72.9  702   9  AV716325   AV716325 AV716325
c 38  606.4  72.6  615   9  AW024324   AW024324 wt69g12.x
  39  600.8  72.0  604  13  BU581717   BU581717 in40c11.y
c 40  597.2  71.5  617  13  BU581396   BU581396 in40c11.x
  41  596    71.4  696  10  BF031281   BF031281 601559030
c 42  592.4  70.9  634  13  BU625384   BU625384 UI-H-PG1-
  43  588.6  70.5  601   9  AV716706   AV716706 AV716706
  44  587.2  70.3  599  14  CB162901   CB162901 K-EST0223
  45  586.6  70.3  662  12  BI862680   BI862680 603389838
```

ALIGNMENTS

```
RESULT 1
AF161398
AF161398
LOCUS       AF161398                845 bp    mRNA    linear   HTC 22-MAY-2001
DEFINITION  Homo sapiens HSPC280 mRNA, partial cds.
ACCESSION   AF161398
VERSION     AF161398.1  GI:6841209
KEYWORDS    HTC.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 845)
AUTHORS     Ye,M., Zhang,Q.H., Zhou,J., Shen,Y., Wu,X.Y., Guan,Z.Q., Wang,L.,
            Fan,H.Y., Mao,Y.F., Dai,M., Huang,Q.H., Chen,S.J. and Chen,Z.
TITLE       Human partial CDS from cd34+ stem cells
JOURNAL     Unpublished
REFERENCE   2  (bases 1 to 845)
AUTHORS     Ye,M., Zhang,Q.H., Zhou,J., Shen,Y., Wu,X.Y., Guan,Z.Q., Wang,L.,
            Fan,H.Y., Mao,Y.F., Dai,M., Huang,Q.H., Chen,S.J. and Chen,Z.
TITLE       Direct Submission
JOURNAL     Submitted (14-MAY-1999) Shanghai Institute of Hematology, Shanghai
            Second Medical University, Rui-Jin Hospital, 197 Rui-Jin Road II,
            Shanghai 200025, P. R. China
FEATURES             Location/Qualifiers
     source          1..845
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
```

RESULT 2
AF116682

LOCUS   AF116682   876 bp   mRNA   linear   HTC 08-MAY-2001
DEFINITION Homo sapiens PRO2013 mRNA, complete cds.
ACCESSION   AF116682
VERSION   AF116682.1 GI:7959862
KEYWORDS   HTC.
SOURCE   Homo sapiens (human)
  ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 876)
  AUTHORS   Zhang,C., Yu,Y., Zhang,S., Wei.H., Zhou,G., Ouyang,S., Luo,L.,
            Bi,J., Liu,M. and He,F.
  TITLE   Functional prediction of the coding sequences of 121 new genes
            deduced by analysis of cDNA clones from human fetal liver
  JOURNAL   Unpublished
REFERENCE   2 (bases 1 to 876)
  AUTHORS   Zhang,C., Yu,Y., Zhang,S., Wei.H., Zhou,G., Ouyang,S., Luo,L.,
            Bi,J., Liu,M. and He,F.
  TITLE   Direct Submission
  JOURNAL   Submitted (24-DEC-1998) Department of Experimental Hematology,
            Institute of Radiation Medicine, Beijing Taiping Road 27, Beijing
            100850, P. R. China
FEATURES   Location/Qualifiers
     source   1..876
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="FLB7547"
                     /tissue_type="liver"
                     /dev_stage="fetal"
     CDS   136..381
                     /note="predicted protein of HQ2013"
                     /codon_start=1
                     /product="PRO2013"
                     /protein_id="AAF71102.1"
                     /db_xref="GI:7959863"
                     /translation="MNVDHEVNLLVEEIHRLGSKNADGKLSVKFGVLFRDDKCANLFE
                     ALVGTLKAAKRRKIVTYPGELLLQGVHDDVDIILLQD"

ORIGIN
Query Match   91.3%; Score 762.6; DB 11; Length 876;
Best Local Similarity 99.5%; Pred. No. 2.7e-171;
Matches 765; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy   67 TCTTTCTTTGCCTCAGCCACTTCCTTCCTTCGCCTCACCTCCCCAGTGCACTGAAGA 126
          |||||||||||||||||||||||||||||||||||||| |||||||||||||||||||
Db   6 TTTTTTTTTTGCCTCAGCCACTTCCTTCCTTCGCCTCACCTCCCCAGTGCACTGAAGA 65

Qy   127 AGGTAACCGGGTCCAGACCCACGCGGCGCCACGCGGCGCCAGTTCTCCGGGGAAAACCGCGCAG 186
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   66 AGGTAACCGGGTCCAGACCCACGCGGCGCCACGCGGCGCCAGTTCTCCGGGGAAAACCGCGCAG 125

Qy   187 AGAGGCAGCAATGAATGTGGATCACGAGGTTAACCTTAGTTGGGTCCTCTTCCGTGATGA 246
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   126 AGAGGCAGCAATGAATGTGGATCACGAGGTTAACCTTAGTTGGGTCCTCTTCCGTGATGA 185

Qy   247 GGGTTCAAAAATGCTGATGAAAGTTAAGCGTAGGAACTGGTCAAAACGAAGGAA 306
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   186 GGGTTCAAAAATGCTGATGAAAGTTAAGCGTAGGAACTGGTCAAAACGAAGGAA 245

Qy   307 TAAATGTGCCAACCTCTTTGAAGCATTGGTGCTTCTGCAAGGTGTTCATGATGATGA 366
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   246 TAAATGTGCCAACCTCTTTGAAGCATTGGTGCTTCTGCAAGGTGTTCATGATGATGA 305

Qy   367 GATTGTAACATATCCAGGAGAGCTGCTTCTGCAAGGTGTTTACATATCTTTATGTACTGCCATTATT 426
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   306 GATTGTAACATATCCAGGAGAGCTGCTTCTGCAAGGTGTTTACATATCTTTATGTACTGCCATTATT 365

Qy   427 ATTACTGCAAGATTAATGTGGTTTACATATCTTTATGTACTGCCATTTTTGTTTCTGGT 486

5'-AAGCAGTGGTATCAACGCAGAGTGGCCATTACGGCCGGG-3' and
5'-ATTCTAGAGGCCGAGCGGCCGACATG-dT(30)NN-3'. Full-length
enriched library was constructed using the Clontech
Creator SMART kit and size-selected to contain the >2 kb
size fraction (other fractions present in NIH_MGC 126 and
NIH_MGC 127). Library created in the laboratory of T.
Usdin, M.D., Ph.D. (NIMH, NIH). Note: this is a NIH_MGC
Library."

ORIGIN

Query Match   89.9%;   Score 750.6; DB 13;   Length 816;
Best Local Similarity   99.3%;   Pred. No. 2e-168;
Matches 764; Conservative 0; Mismatches 4; Indels 1; Gaps 1;

```
Qy   60 ACAGCCGTCTTTCTTTGCCTCAGCCACTTCCTTCCTTCGCCTCACCTCCCCCAGTGCA 119
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4 ACAGCCGTCTTTCTTTGCCTCAGCCACTTCCTTCCTTCGCCTCACCTCCCCCAGTGCA 63

Qy  120 CTGAAGAGGTAACCGGTCCAGACCCACGGCGCGCAGTTCTCCGGCGGGAAGGAAAAC 179
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   64 CTGAAGAGGTAACCGGTCCAGACCCACGGCGCGCAGTTCTCCGGCGGGAAGGAAAAC 123

Qy  180 CGGCCAGAGAGGCAGCAATGAAGTGGATCACGAGGTAACCTCTTAGTGGAGGAAATTC 239
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  124 CGGCCAGAGAGGCAGCAATGAAGTGGATCACGAGGTAACCTCTTAGTGGAGGAAATTC 183

Qy  240 ATGTTTTGGGTTCAAAAAATGCTGATGGAAAGTTAAGCGTGAAAATTTGGGGTCCTCTTCC 299
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  184 ATCTTTGGTTCAAAAAATGCTGATGGAAAGTTAAGCGTGAAAATTTGGGGTCCTCTTCC 243

Qy  300 GTGATGATAAATGTGCCAACCTCTTTGAAGCATTGGTAGGAACTCTTAAAGTCTGCAAAAC 359
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  244 GTGATGATAAATGTGCCAACCTCTTTGAAGCATTGGTAGGAACTCTTAAAGTCTGCAAAAC 303

Qy  360 GAAGGAAGATTGTAACATATCCAGGAGAGCTGCTTCGCAAGGTGTTCATGATGATGTTG 419
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  304 GAAGGAAGATTGTAACATATCCAGGAGAGCTGCTTCGCAAGGTGTTCATGATGATGTTG 363

Qy  420 ACATTATATTACTGCAAGATGTGTTTACATATCTTTATGTACTGCCATTTTTGT 479
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  364 ACATTATATTACTGCAAGATAATGTGGTTTACATATCTTTATGTACTGCCATTTTTTGT 423

Qy  480 TTCTGGTAAACTGGAATATAAAGTGAAAGAACAAACATTTGAACATTTGAACATTTAATGTATTT 539
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  424 TTCTGGTAAACTGGAATATAAAGTGAAAGAACAAACATTTGAACATTTAATGTATTT 483

Qy  540 TATAGAACTTTGTAAACGAAAGGAGATTCATGTTTTAGAGAGTCTGTCCTTTTTATATCT 599
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  484 TATAGAACTTTGTAAACGAAAGGAGATTCATGTTTTAGAGAGTCTGTCCTTTTTATATCT 543

Qy  600 TGAAAGAAAATCATGTATGATGCTATAAAATAAAATAAAATCCTATTATTTTCTCAGGAATCTG 659
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  544 TGAAAGAAAATCATGTATGATGCTATAAAATAAAATAAAATCCTATTATTTTCTCAGGAATCTG 603

Qy  660 GTTAGGAATTGCAGGCAATGACAGAGATTTTTGCGGGATGGGAATGTTTGTTCATAA 719
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  604 GTTAGGAATTGCAGGCAATGACAGAGATTTTTGCGGGCAGGGATGGGAATGTTTGTTCATAA 663

Qy  720 ATAATTAGACATTTTCTATAGATATTTGACATTCGCGAAAGCACAAGCAACAAACTGAAGA 779
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  664 ATAATTAGACATTTTCTATAGATATTTGACATTCGCGAAAAGCAC-AGCAAACTGAAGA 722

Qy  780 CCAACTCCTATGAGAAATATTATATGATGTTTATGTAATAAAGACATGTAA 828
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  723 CCAACTCCTATGAGAAATATTATATGATGTTTATGTAATAAAGACATGTTA 771
```

RESULT 4
BI859694
LOCUS   BI859694   797 bp   mRNA   linear   EST 10-OCT-2001
DEFINITION   603386283F1 NIH_MGC_87 Homo sapiens cDNA clone IMAGE:5394836 5',
             mRNA sequence.
ACCESSION   BI859694

```
366 ATTACTGCAAGATAATGTGGTTACATATCTTTATGTACTGCCATTTTTGTTCTGGT 425
     ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
487 AAACTGGAATATAAAGTGAAGAACAAACATTTGAACATACTTAATGTATTTTTATAGAA 546
     |||||||||||||||||||||||||||||||||||||||||||||||||||||
426 AAACTGGAATATAAAGTGAAGAACAAACATTTGAACATACTTAATGTATTTTTATAGAA 485

547 CTTTGTAAACGAAAGGAAGGAGATTCATGTTTTAGAGAGTCTGTCCTTTTTTATATCTTGAAAGA 606
     ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
486 CTTTGTAAACGAAAGGAAGGAGATTCATGTTTTAGAGAGTCTGTCCTTTTTTATATCTTGAAAGA 545

607 AAATCATGTATGATGCTATAAAATAAAATCCATTATTTTTCTCAGGAATCCTGGTTAGGA 666
     ||||||||||||||||||||||||||||||||||||||||||||||||||||||
546 AAATCTATGTATGATGCTATAAAATCCTATTATTTTTCTCAGGAATCCTGGTTAGGA 605

667 ATTGCAGGCAATGACAATAGAGATTTTTGCGGGGCAGGGATGGGAATGTTTGTTCATAAAATAATTA 726
     |||||||||||||||||||||||||||||||||||||||||||||||||||||
606 ATTGCAGGCAATGACAATAGAGATTTTTGCGGGGCAGGGATGGGAATGTTTGTTCATAAAATAATTA 665

727 GACATTTTCTATAGATATTGACATTCTGCGAAAGCAACAAGCAAACTGAAGACCAACTC 786
     ||||||||||||||||||||||||||||||||||||||||||||||| |||||
666 GACATTTTCTATAGATATTGACATTCTGCGAAAGCAACAAGCAAACTGAAGACCAACTC 725

787 CTATGAGAAATATTATGATGTTTATGTAATAAAGACATGTAACTGTCTT 835
     ||||||||||||||||||||||||||||||||||||||||||||||||
726 CTATGAGAAATATTATGATGTTTATGTAATAAAGACATGTAACTGTCTT 774
```

JULT 3
349360
CUS   BU849360   816 bp   mRNA   linear EST 16-OCT-2002
FINITION AGENCOURT 10432847 NIH_MGC_128 Homo sapiens cDNA clone
         IMAGE:6597818 5', mRNA sequence.
CESSION BU849360
RSION   BU849360.1 GI:24034323
YWORDS   EST.
JRCE   Homo sapiens (human)
ORGANISM Homo sapiens
         Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
         Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
FERENCE 1 (bases 1 to 816)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE   National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
MMENT   Contact: Robert Strausberg, Ph.D.
         Email: cgapbs-r@mail.nih.gov
         Tissue Procurement: NCI
         cDNA Library Preparation: Michael Brownstein Laboratory
         cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
         DNA Sequencing by: Agencourt Bioscience Corporation
         Clone distribution: MGC clone distribution information can be
         found through the I.M.A.G.E. Consortium/LLNL at:
         http://image.llnl.gov
         Plate: LLCM2829 row: k column: 02
         High quality sequence stop: 604.
ATURES   Location/Qualifiers
     source   1..816
                   /organism="Homo sapiens"
                   /mol_type="mRNA"
                   /db_xref="taxon:9606"
                   /clone="IMAGE:6597818"
                   /tissue_type="mixed (pool of 40 RNAs)"
                   /lab_host="DH10B (Ti-phage-resistant)"
                   /clone_lib="NIH_MGC_128"
                   /note="Vector: pDNR-LIB; Site 1: SfiI (ggccattatggcc);
                   Site 2: SfiI (ggccagcctcggcc); Double-stranded cDNA was
                   prepared from a pool of 40 cell line polyA+ RNAs (bladder
                   - 2%, blood - 33.4%, brain - 5.6%, breast - 12.5%, colon -
                   4%, connective tissue - 1.4%, eye - 1%, intestine - 2.6%,
                   kidney - 2.2%, liver - 5.7%, lung - 10.8%, NK-cell -
                   5.2%, ovary - 4%, pharynx - 2.5%, prostate - 4.3%,
                   salivary gland - 1.3%, and skin - 2.3%. 5' and 3'
                   adaptors were used in cloning as follows:
```

```
RSION    BI859694.1  GI:16000441
YWORDS   EST.
URCE     Homo sapiens (human)
ORGANISM Homo sapiens
         Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
         Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
FERENCE  1  (bases 1 to 797)
AUTHORS  NIH-MGC http://mgc.nci.nih.gov/.
TITLE    National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL  Unpublished (1999)
MMENT    Contact: Robert Strausberg, Ph.D.
         Email: cgapbs-r@mail.nih.gov
         Tissue Procurement: DCTD/DTP
         cDNA Library Preparation: Life Technologies, Inc.
         cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
         DNA Sequencing by: Incyte Genomics, Inc.
         Clone distribution: MGC clone distribution information can be
         found through the I.M.A.G.E. Consortium/LLNL at:
         http://image.llnl.gov
         Plate: LLAM12006 row: f  column: 21
         High quality sequence stop: 787.
ATURES   Location/Qualifiers
source   1..797
         /organism="Homo sapiens"
         /mol_type="mRNA"
         /db_xref="taxon:9606"
         /clone="IMAGE:5394836"
         /tissue_type="mammary adenocarcinoma, cell line"
         /lab_host="DH10B (phage-resistant)"
         /clone_lib="NIH_MGC 87"
         /note="Organ: Breast; Vector: pCMV-SPORT6; Site_1: NotI;
         Site_2: SalI; Cloned unidirectionally; oligo-dT primed;
         Average insert size 1.383 kb. Library enriched for
         full-length clones and constructed by Life Technologies.
         Note: this is a NIH_MGC Library."
IGIN

Query Match      88.3%;  Score 737;  DB 12;  Length 797;
Best Local Similarity   97.7%;  Pred. No. 3.5e-165;
Matches  779;  Conservative   0;  Mismatches  15;  Indels   3;  Gaps   3;

/     20 GAAACACTACTTTTCCCCTACCCTGCTCCTCCTCCTCCCACAGCCGTCTTCTTTCTTTGC   79
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
o      1 GAAACACTACTTTTCCCCTACCCTGCTCCTCCTCCTCCCACAGCCGTCTTCTTTCTTTGC   60

/     80 CTCAGCCACTTCCTTCCTTCCTTCGCCTCACCCTCCCAGTGCACTGAAGAAGGTAACCGGGTC  139
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
o     61 CTCAGCCACTTCCTTCCTTCCTTCGCCTCACCCTCCCAGTGCACTGAAGAAGGTAACCGGGTC  120

y    140 CAGACCCACGGCGGCGGCCAGTTCTCCGGCGGAAGGAAAACCGGCGAGAGAGGCAGCAATG  199
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b    121 CAGACCCACGGCGGCGGCCAGTTCTCCGGCGGAAGGAAAACCGGCGAGAGAGGCAGCAATG  180

y    200 AATGTGGATCACGAGGTTAACCTCTTAGTGGAGGAAAATTCATCGTTCGTTCAAAAAAT  259
         |||||||||||||||||||||||||||||||||||||||||||| |||||||||||||
b    181 AATGTGGATCACGAGGTTAACCTCTTAGTGGAGGAAAATTCATCGTTCGTTGGTTCAAAAAAT  240

y    260 GCTGATGGAAAGTTAAGCGTGAAATTGGGGGTCCTTCTTCCGTGATGAAATGTGCCAAC  319
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b    241 GCTGATGGAAAGTTAAGCGTGAAATTGGGGGTCCTTCTTCCGTGATGATGAAATGTGCCAAC  300

y    320 CTCTTTGAAGCATTGGTAGGAAACTCTTAAAGCTGCAAAACGAAGGAAGATTGTAACATAT  379
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b    301 CTCTTTGAAGCATTGGTAGGAAACTCTTAAAGCTGCAAAACGAAGGAAGATTGTAACATAT  360

y    380 CCAGGAGGACTGCTTCTGCAAGGTGTTCATGATGTTGACATTATATTACTGCAAGAT  439
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b    361 CCAGGAGGACTGCTTCTGCAAGGTGTTCATGATGTTGACATTATATTACTGCAAGAT  420

y    440 TAATGTGGTTTACATATCTTTATGTACTGCCATTTTTGTTTCTGGTAAACTGGAATATA  499
         ||||||||||||||||||||||||||||||||||||     |||||||||||||||||
b    421 TAATGTGGTTTACATATCTTTATGTACTGCCA-TTTTTGTTTCTGGTAAACTGGAATATA  479
```

```
500 AAGTGAAAGAACAAACATTTGAACATACTTAATGTATTTTATAGAACTTTGTAAACGAA  559
Qy  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  480 AAGTGAAAGAACAAACATTTGAACATACTTAATGTATTTTATAGAACTTGGTAAACGAA  539

Qy  560 AGGAGATTCATGTTTAGAAGTCTGTCCTTTTTATATCTTGAAAGAAAATCATGTATG  619
    |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  540 AGGAGATTCATGTTTAGAAGTCTGTCCTTTTTATATCTTGAAAGAAAATCATGTATG  599

Qy  620 ATGCTAT-AAAATAAATCCTATTATTTTCTCAGGAATCTGGTTAGGAATTGCAGGCAAT  678
    ||||||| |||||||||||||||||||||||||||||||||||||||||||||||||
Db  600 ATGCTATAAAAATAAATCCTATTATTTTTCTCAGGAATCTGGTTAGGAATTGCAGGCAAT  659

Qy  679 GAGATTTTTTGCGGGGCAGGATGGGAATGTTGTTCATAA-ATAATTAGACATTTTCTA  737
    |||||||||||||||||||||||||||||||||||||| |||||||||||||||||||
Db  660 GAGATTTTTTGCGGGGCAGGGATGGGAATGTTCGTTCATAAACATAATTAGACATTTTCTA  719

Qy  738 TAGATATTTGACATTCTGCGAAAGCAACAAGCAACAAGCAACCAACTCCTATGAGAAAT  797
    |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  720 TAGATATTTGACATTCTGCGAAAGCAACAAGCAACAAGCAACCAATCCTATGAGGAAAT  779

Qy  798 ATTATGATGTTTATGTA  814
    |||||||||||||||||
Db  780 ATTATGCATGTTATGTA  796
```

```
RESULT 5
BQ015110/c
LOCUS       BQ015110               749 bp  mRNA    linear   EST  26-MAR-2002
DEFINITION  UI-H-ED1-axw-a-08-0-UI.s1 NCI_CGAP_ED1 Homo sapiens cDNA clone
            IMAGE:5834383 3', mRNA sequence.
ACCESSION   BQ015110
VERSION     BQ015110.1  GI:1974011
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 749)
AUTHORS     NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE       National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
JOURNAL     Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Dr. Jose Mercuende
            cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
            cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
            DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
            Clone Distribution: Clone distribution information can be found
            through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
            The following repetitive elements were found in this cDNA
            sequence: 1-26, >AT_rich#Low_complexity (matched compliment)
            Seq primer: M13 FORWARD
            POLYA=Yes.
FEATURES    Location/Qualifiers
source      1..749
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone="IMAGE:5834383"
            /tissue_type="Chondrosarcoma"
            /dev_stage="Adult"
            /lab_host="DH10B (Life Technologies)"
            /clone_lib="NCI_CGAP_ED1"
            /note="Organ: Left Pubic Bone; Vector: pT7T3-Pac
            (Pharmacia) with a modified polylinker; Site 1: EcoR I;
            Site 2: Not I; NCI_CGAP_ED1 is a normalized cDNA library
            containing the following tissue(s): Chondrosarcoma cell
            line CS5. The library was constructed according to
            Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
            1996. First strand cDNA synthesis was primed with an
            oligo-dT primer containing a Not I site. Double stranded
            cDNA was ligated to an EcoR I adaptor, digested with Not
```

```
KEYWORDS    EST. .
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 741)
AUTHORS     Xu,X., Gu,J., Liu,F., Qu,J., Zhao,M., Li,Y., Huang,Q., Zhou,J.,
            Song,H., Gu,Y., Gu,S., Yang,Y., Gao,G., Xiao,H., Li,N., Qian,B., Gao,X.,
            Cheng,Z., Xu,S., Gu,W., Tu,Y., Jia,J., Fu,G., Ren,S., Zhong,M.,
            Lu,G., Cheng,Z. and Han,Z.
TITLE       Homo sapiens cDNA DCB clones
JOURNAL     Unpublished (2000)
COMMENT     Contact: Zeguang Han
            Chinese National Human Genome Center at Shanghai
            351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
            201203, P. R. China
            Tel: 86-21-50801919(ex.45)
            Fax: 86-21-50801922
            Email: hanzg@chgc.sh.cn
            This clone is available at CHGC in Shanghai.
FEATURES             Location/Qualifiers
     source          1..741
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="DCBBBD06"
                     /cell_type="dendritic cells"
                     /dev_stage="mature"
                     /lab_host="BM25.8"
                     /clone_lib="DCB"
                     /note="Vector: pTriplEx2; Site_1: sfiIA; Site_2: sfiIB"
ORIGIN

Query Match           83.3%;   Score 695.2;  DB 9;   Length 741;
Best Local Similarity 98.6%;   Pred. No. 3.3e-155;
Matches 700; Conservative  0; Mismatches 10; Indels 0; Gaps 0;

Qy    126 AAGGTAACCGGGTCCAGACCCACCGGCGGCCAGTTCTCCGGCGGGAAGGAAACCGCGCA 185
          ||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db      4 AAGGTAACCGGGTCCACACCCACCGGCGGCCAGTTCTCCGGCGGGAAGGAAACCGCGCA 63

Qy    186 GAGAGGCAGCAATGAATGTGGATCACGAGGTTAACCTCTTAGTGGAGGAAATTCATCGTT 245
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     64 GAGAGGCAGCAATGAATGTGGATCACGAGGTTAACCTCTTAGTGGAGGAAATTCATCGTT 123

Qy    246 TGGGTTCAAAAAAATGCTGATGGAAAGTTAAGCGTGAAATTGGGGTCCTTCCGTGATG 305
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    124 TGGGTTCAAAAAAATGCTGATGGAAAGTTAAGCGTGAAATTGGGGTCCTTCCGTGATG 183

Qy    306 ATAAAATGCCAACCTCTTGAAGCATTGGTAGGAACTCCTAAAGCTGCAAAACGAAGGA 365
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    184 ATAAAATGCCAACCTCTTGAAGCATTGGTAGGAACTCCTAAAGCTGCAAAACGAAGGA 243

Qy    366 AGATTGTAACATATCCAGGAGAGCTGCTTCTGCAAGGTGGTTCATGATGATGTTGACATTA 425
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    244 AGATTGTAACATATCCAGGAGAGCTGCTTCTGCAAGGTGGTTCATGATGATGTTGACATTA 303

Qy    426 TATTACTGCAAGATTAATGTGGTTTACATATCTTTATGTGAACATTGAACATACTACTTAATGA 485
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    304 TATTACTGCAAGATTAATGTGGTTTACATATCTTTATGTGAACATTGAACATACTACTTAATGA 363

Qy    486 TAAACTGGAATATAAAGTGAAAGAAGAACAAACATTTGAACATTAATGTATTTTTATAGA 545
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    364 TAAACTGGAATATAAAGTGAAAGAAGAACAAACATTTGAACATTAATGTATTTTTATAGA 423

Qy    546 ACTTTGTAAACGAAAGGAGATTCATGTTTTAGAGTCTGTCCTTTTTTATATCCTTTATATCTTGAAAG 605
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    424 ACTTTGTAAACGAAAGGAGATTCATGTTTTAGAGTCTGTCCTTTTTTATATCCTTTATATCTTGAAAG 483

Qy    606 AAAATCATGTATGATGCTATAAAAATAAATCCTATTATTTTTCTCCAGGAATCGTGGTTAGG 665
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    484 AAAATCATGTATGATGCTATAAAAATAAATCCTATTATTTTTCTCCAGGAATCGTGGTTAGG 543
```

```
I, and cloned directionally into pT7T3-Pac vector. The
oligonucleotide used to prime the synthesis of
first-strand cDNA contains a library tag sequence that is
located between the Not I site and the (dT)18 tail. The
sequence tag for this library is GCTCAAGGCT.
TAG_LIB=UI-H-ED1
TAG_SEQ=GCTCAAGGCT"
GIN

uery Match           85.7%;   Score 715.6;  DB 12;  Length 749;
est Local Similarity 99.0%;   Pred. No. 4.4e-160;
atches 718; Conservative  0; Mismatches  7; Indels 0; Gaps 0;

111 CCCAGTGCACTGAAGAGAGGTAACCGGGTCCAGACCCACCGGCGGCCAGTTCTCCGGCGGG 170
    ||||||||||||||||||| ||||||||||||||||||||||||||||||||||| | |||
749 CCCAGTGCACTGAAGAGAGTNAACCGGGTCCAGACCCACCGGCGGCCAGTTCTCCGCGNGG 690

171 AAGGAAACCGCGCAGAGAGGCAGCAATGAATGTGGATCACGAGGTTAACCTCTTAGTTGG 230
    |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
689 AAGGAAACCGCGCAGAGAGGCAGCAATGAATGTGGATCACGAGGTTAACCTCTTAGTTGG 630

231 AGGAAATTCATCGTTTGGGTTCAAAAAAATGCTGATGGAAAGTTAAGCGTGAAATTTGGGG 290
    |||||||||||||| |||||||||||||||||||||||||||||||||||||||||||||
629 AGGAAATTCACCGTNTGGGTTCAAAAAAATGCTGATGGAAAGTTAAGCGTGAAATTTGGGG 570

291 TCCTCTTCCGTGATGATAAAATGCCAACCTCTTTGAAGCATTGGTAGGAACTCTTAAAG 350
    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
569 TCCTCTTCCGTGATGATAAAATGCCAACCTCTTTGAAGCATTGGTAGGAACTCTTAAAG 510

351 CTGCAAAACGAAGCAAGATTGTAACATATCCAGGAGAGAGCTGCTTCTGCAAGGTGTTCATG 410
    |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
509 CTGCAAAACGAAGCAAGATTGTAACATATCCAGGAGAGAGCTGCTTCTGCAAGGTGTTCATG 450

411 ATGATGTTGACATTATTATTACTGCAAGATTAATGTGGTTTACATATCTTTATGTACTGCC 470
    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
449 ATGATGTTGACATTATTATTACTGCAAGATTAATGTGGTTTACATATCTTTATGTACTGCC 390

471 ATTTTTTGTTTCTCGGTAAACTGGAATATAAAGTGAAAGAAGAACAAACATTTGAACATTA 530
    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
389 ATTTTTTGTTTCTCGGTAAACTGGAATATAAAGTGAAAGAAGAACAAACATTTGAACATTA 330

531 ATGTATTTTTATAGAACTTTGTAAACGAAGGAGATTCATGTTTTAGAGTCTGTCCTTT 590
    |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
329 ATGTATTTTTATAGAACTTTGTAAACGAAGGAGATTCATGTTTTAGAGTCTGTCCTTT 270

591 TTTATATCTTGAAAGAAAAATCATGTATGATGCCTATTAAAATCCTATTATTTTTCTC 650
    |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
269 TTTATATCTTGAAAGAAAAATCATGTATGATGCCTATTAAAATCCTATTATTTTTCTC 210

651 AGGAATCTGGTAGGAATTGCAGGGCAATGAGAGATTTTTGCGGGGATGGGAATGTT 710
    |||||||||| |||||||||||||||||||||||||||||||||||||||||||||
209 AGGAATCTGGTTAGGAATTGCAGGGCAATGAGAGATTTTTGCGGGGATGGGAATGTT 150

711 TGTTCATAAAATAATTAGACATTTTCATAGAGAATATTTGACATTCTGCGAAAGCAACAAGCA 770
    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
149 TGTTCATAAAATAATTAGACATTTTCATAGAGAATATTTGACATTCTGCGAAAGCAACAAGCA 90

771 AACTGAAGACCAACTCCTATGAGAAATAATATGATGTTTATGTGTAATAAGACATGTAACT 830
    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
89 AACTGAAGACCAACTCCTATGAGAAATAATATGATGTTTATGTGTAATAAGACATGTAACT 30

831 GTCTT 835
    |||||
29 GTCTT 25
```

```
SULT 6
717724
CUS       AV717724    741 bp    mRNA    linear    EST 16-OCT-2000
FINITION  AV717724 DCB Homo sapiens cDNA clone DCBBBD06 5', mRNA sequence.
CESSION   AV717724
RSION     AV717724.1  GI:10814876
```

```
  666 AATTGCAGGCAATGAGAGATTTTTGCGGGGCAGGGATGGGAATGTTTGTTCATAAATAATT 725
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||
  544 AATTGCAGGCAATGAGAGATTTTTGCGGGGCAGGGATGGGAATGTTTGTTCATAAATAATT 603

  726 AGACATTTCTATAGATATTTGACATTCTCGAAGACAAGCAACTGAAGACCAACT 785
      |||||||||||||||||||||||||||  ||||||    ||||||||||||||||
  604 AGACATTTCTATAGATATTTGACATTCTCGAAAGNCACAAGCAAACTGAAGACCCCAAC 663

  786 CCTATGAGAAATATTATGATGTTTATGTAATAAAGACATGTAACTGTCTT 835
      |||||||||||||||||||||||||| |||||||||||||||||||||||
  664 TCTATGAGAAATATTATGATGTTTATNGTATAAAGACATGTAACTGTCTT 713

SULT 7
716363

CUS         AV716363                752 bp    mRNA    linear   EST 11-OCT-2000
FINITION    AV716363 DCB Homo sapiens cDNA clone DCBARC08 5', mRNA sequence.
CESSION     AV716363
RSION       AV716363.1  GI:10797880
YWORDS      EST.
URCE        Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
FERENCE     1  (bases 1 to 752)
AUTHORS     Xu,X., Gu,J., Liu,F., Qu,J., Zhao,M., Li,Y., Huang,Q., Zhou,J.,
            Song,H., Gu,Y., Yang,Y., Gao,G., Xiao,H., Li,N., Qian,B., Gao,X.,
            Cheng,Z., Xu,S., Gu,W., Tu,Y., Jia,J., Fu,G., Ren,S., Zhong,M.,
            Lu,G., Cheng,Z. and Han,Z.
TITLE       Homo sapiens cDNA DCB clones
JOURNAL     Unpublished (2000)
MMENT       Contact: Zeguang Han
            Chinese National Human Genome Center at Shanghai
            351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
            201203, P. R. China
            Tel: 86-21-50801919(ex.45)
            Fax: 86-21-50801922
            Email: hanzg@chgc.sh.cn
            This clone is available at CHGC in Shanghai.
EATURES     Location/Qualifiers
source      1..752
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone="DCBARC08"
            /cell_type="dendritic cells"
            /dev_stage="mature"
            /lab_host="BM25.8"
            /clone_lib="DCB"
            /note="vector: pTriplEx2; Site_1: sfiIA; Site_2: sfiIB"

RIGIN

Query Match      83.2%;    Score 694.6;  DB 9;   Length 752;
Best Local Similarity 97.4%;  Pred. No. 4.6e-155;
Matches 703;   Conservative   0;   Mismatches  19;   Indels   0;   Gaps   0;

 y  114 AGTGCACTGAAGAAGGTAACCGGTCCAGACCTCCAGTTCTCCGGCGGGGAAG 173
         |||||||||||||||||||||||||||||||||||||||||||||||||||||
 b    1 AGTGCACTGAAGAAGGTAACCGGGTCCAGACCCCAGTTCTCCGGCGGGGAAG 60

 y  174 GAAAAACCGCGCAGAGAGGCAGCAATGAATGTGGATCACGAGGTTAACCTCT 233
         |||||||||||||||||||||||||||||||||||||||||||||||||||||
 b   61 GAAAAACCGCGCAGAGAGGCAGCAATGAATGTGGATCACGAGGTTAACCTCTTAGTGGAGG 120

 y  234 AAATTCATCGTTGGGTTCAAAAAATGCTGATGATGGAAAGTTAAGCGTGAAATTTGGGGTCC 293
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
 b  121 AAATTCATCGTTGGGTTCAAAAAATGCTGATGATGGAAAGTTAAGCGTGAAATTTGGGGTCC 180

 y  294 TCTTCCGTGATGATAAATGTGCCAACCTCTTGAAGCATTGGTAGGAACTCTTAAAGCTG 353
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
 b  181 TCTTCCGTGATGATAAATGTGCCAACCTCTTTGAAGCATTGGTAGGAACTCTTAAAGCTG 240

 y  354 CAAAACGAAGGAAGATTGTAACATATCCAGGAGAGAGCTGCTTCTCGCAAGGTGTTCATGATG 413
```

```
  241 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
  241 CAAAACGAAGGAAGATTGTAACATATCCAGGAGAGGCTGCTTCTCGCAAGGTGTTCATGATG 300

  414 ATGTTGACATTATTATTACTGCAAGATTAATGTGTTACATATCTTATGTACTGCCATT 473
      |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
  301 ATGTTGACATTATTATTACTGCAAGATTAATGTGTTACATATCTTATGTACTGCCATT 360

  474 TTTGTTCTGTGTAAACTGAAATAAAGTGAAAGAACAAACATTTGAACATACTTAATG 533
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
  361 TTTGTTCTGTGTAAACTGAAATAAAGTGAAAGAACAAACATTTGAACATACTTAATG 420

  534 TATTTTTATAGAACTTTGTAAACGAAAGGAGATTCATGTTTTAGAAGTCTGTCCTTTTTT 593
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
  421 TATTTTTATAGAACTTTGTAAACGAAAGGAGATTCATGTTTTAGAAGTCTGTCCTTTTTT 480

  594 ATATCTTGAAAGAAAATCATGTATGATGCTATAAAATAAATCCTATATATTTTTCTCAGG 653
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
  481 ATATCTTGAAAGAAAATCATGTATGATGCTATAAAATAAATCCTATATATTTTTCTCAGG 540

  654 AATCTGGTTAGGAATTGCAGGCAATGACAGATTTTTTGCGGGGCAGGGATGGGAATGTTTGT 713
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
  541 AATCTGGTTAGGAATTGCAGGCAATGACAGATTTTTTGCGGGGCAGGGATGGGAATGTTTGT 600

  714 TCATAAATTATTAGACATTTCTATAGATATTTGACATTCTGCGAAAGCAACAAGCAAAC 773
      ||||||||||||||||||||||||||||  |||||||||| ||||||||||  |||||||
  601 TCATAAATTATTAGACATTTCTATAGATATTTGACATTCTNGCGAAAGCACANGCANAC 660

  774 TGAAGACCAACTCCTATAGAGAAATATTATGATGTTTATGTAATAAAGACATGTAACTGTC 833
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
  661 TGAAGACCAACTCCTATAGAGAAATATTATGATGTTTATGTAATAAAGACATGTAACTGTC 720

  834 TT 835
      ||
  721 TT 722
```

```
RESULT 8
AV713584
LOCUS       AV713584                714 bp    mRNA    linear   EST 11-OCT-2000
DEFINITION  AV713584 DCB Homo sapiens cDNA clone DCBBBB10 5', mRNA sequence.
ACCESSION   AV713584
VERSION     AV713584.1  GI:10795101
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 714).
AUTHORS     Xu,X., Gu,J., Liu,F., Qu,J., Zhao,M., Li,Y., Huang,Q., Zhou,J.,
            Song,H., Gu,Y., Yang,Y., Gao,G., Xiao,H., Li,N., Qian,B., Gao,X.,
            Cheng,Z., Xu,S., Gu,W., Tu,Y., Jia,J., Fu,G., Ren,S., Zhong,M.,
            Lu,G., Cheng,Z. and Han,Z.
TITLE       Homo sapiens cDNA DCB clones
JOURNAL     Unpublished (2000)
COMMENT     Contact: Zeguang Han
            Chinese National Human Genome Center at Shanghai
            351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
            201203, P. R. China
            Tel: 86-21-50801919(ex.45)
            Fax: 86-21-50801922
            Email: hanzg@chgc.sh.cn
            This clone is available at CHGC in Shanghai.
FEATURES    Location/Qualifiers
source      1..714
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone="DCBBBB10"
            /cell_type="dendritic cells"
            /dev_stage="mature"
            /lab_host="BM25.8"
            /clone_lib="DCB"
            /note="vector: pTriplEx2; Site_1: sfiIA; Site_2: sfiIB"
```

```
GIN

Query Match    81.9%;   Score 683.8; DB 9; Length 714;
Best Local Similarity 98.6%;  Pred. No. 1.7e-152;
Matches 699;  Conservative  0;  Mismatches  0;  Indels  1;  Gaps  1;

128 GGTAACCGGGTCCAGACCCACGGGCGCCAGTTCTCCGGCGGGAAGGAAAACCGCGCAGA 187
    |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
  1 GGTAACCGGGTCCAGACCCACGGGCGCCAGTTCTCCGGCGGGAAGGAAAACCGCGCAGA 60

188 GAGGGCAGCAATGAATGTGATCACGAGGTTAACCTCTTAGTGGAGGAAATTCATCGTTTG 247
    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 61 GAGGGCAGCAATGAATGTGATCACGAGGTTAACCTCTTAGTGGAGGAAATTCATCGTTTG 120

248 GGTTCAAAAAATGCTGATGGAAAGTTAAGCGTGGAAATTGGGGTCCTCTTCCGTGATGAT 307
    |||||||||||||||||||||||||||||||||||||||||||| |||||||||||||||
121 GGTTCAAAAAATGCTGATGGAAAGTTAAGCGTGGAAATTGTGGTCCTCTTCCGTGATGAT 180

308 AAAATGTGCCAACCTCTTGAAGCATTGGTAGGAGAACTCTAAAGCTGCAAAACGAAGGAAG 367
    |||||||||||||||||||||||||||||||||||||||||||||||||||||||
181 AAAATGTGCCAACCTCTTGAAGCATTGGTAGGAGAACTCTAAAGCTGCAAAACGAAGGAAG 240

368 ATTGTAACATATCCAGGAGAGCTGCTTCGCAAGTGTTCATGATGATGTTGACATTATA 427
    |||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
241 ATTGTAACATATCCAGGAGAGCTGCTTCTGCAAGTGTTCATGATGATGTTGACATTATA 300

428 TTACTGCAAGATTAATGTGGTTTACATATCTTTATGTACTGCCATTTTTGTTTCTGGTA 487
    ||||||||||||||||||||||||||||||||||||||||||| |||||||||||||
301 TTACTGCAAGATTAATGTGGTTTACATATCTTTATGTACTGCCATTTTTTGTTTCTGGTA 360

488 AACTGGAATATAAAGTGAAAGAACAAACATTTGAACATACTTAATGTATTTTTATAGAAC 547
    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
361 AACTGGAATATAAAGTGAAAGAACAAACATTTGAACATACTTAATGTATTTTTATAGAAC 420

548 TTTGTAAACGAAAGGAGATTCATGTTTTAGAAGTCTGTCCTTTTTTTATATCTTGAAAGAA 607
    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
421 TTTGTAAACGAAAGGAGATTCATGTTTTAGAAGTCTGTCCTTTTTTTATATCTTGAAAGAA 480

608 AATCTATGTATGATGCTATAAAATAAATCCTATATATTTTTCTCAGGAATCTGGTTAGGAA 667
    |||||||||||||||||||||||||||||||||||||||||||||||||||||||
481 AATCTATGTATGATGCTATAAAATAAATCCTATATATTTTTCTCAGGAATCTGGTTAGGAA 540

668 TTGCAGGCAATGAGATTTTTTGCGGCGCAGGGATGGGAATGTTTTGTTCATAAAATAATTAG 727
    ||||||||||||||||||||| |||||||||| |||||||||||||||||||||||||
541 TTGCAGGCAATGAGATTTTTTNGCGGCGCAGGGGTGGGAATGTTTTGTTCATAAAATAATTAG 600

728 ACATTTTCTATAGATATTTGACATTCTGCGAAAGCAACA-GCAAACTGAAGACCAACTC 786
    ||||||||||||||||||||||||||||||||||||| |||||||||||||||
601 ACATTTTCTATAGATATTTGACATTCTGCGAAAGCAACAAANGCAACCTGAAGACCAACTC 660

787 CTATGAGAAAATATTATGATGTTTATGTAATAAAGACATGTAACTGCTT 835
    |||||||||||||||||||||||||||||| ||||||| |||
661 CTATGAGAAAATATTATGATGTTTATGTAATAAAGAACATGNTACTGTCTT 709
```

National Cancer Institute / NIH
Bldg. 31 Rm10A07 Bethesda, MD 20892
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: CLONTECH Laboratories, Inc.
cDNA Library Preparation: Narayan Bhat
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: NDCM209 row: a column: 24
High quality sequence stop: 543.

```
FEATURES          Location/Qualifiers
    source        1..906
                  /organism="Homo sapiens"
                  /mol_type="mRNA"
                  /db_xref="taxon:9606"
                  /clone="IMAGE:30413687"
                  /tissue_type="Pooled"
                  /lab_host="DH10B (T1 phage-resistant)"
                  /clone_lib="NIH_MGC_191"
                  /note="Vector: pDNR-LIB; Site 1: SfiI (ggccattatggcc);
                  Site 2: SfiI (ggccgcctcggcc); Library is oligo-dT primed
                  and directionally cloned. PBMC - Peripheral Blood
                  Mononuclear Cells. RNA was pooled from 3/6hour stimulation
                  with PMA adn Ionomycin. 5' and 3' adaptors were used in
                  cloning as follows: 5' adaptor sequence:
                  5'-CACGGCCATTATGGCC-3' and 3' adaptor sequence:
                  5'-ATTCTAGAGGCCGAGGCGGCCGACATG-dT(30)BN-3' (where B = A,
                  C, or G and N = A, C, G, or T). Average insert size 1.69
                  kb (range 0.70-5.0 kb). 15/15 colonies contained inserts
                  by PCR. This library was enriched for full-length clones
                  and was constructed by Clontech Laboratories (Palo Alto,
                  CA). Note: this is a NIH_MGC Library."

ORIGIN

Query Match    80.7%;   Score 673.8; DB 14;  Length 906;
Best Local Similarity 96.0%;  Pred. No. 4.4e-150;
Matches 713;  Conservative  0;  Mismatches  27;  Indels  3;  Gaps  2;

Qy  59 CACAGCCGTCTTCTCTTTGCCTCAGCCACTCCTTCCTTCCGCCTCACCTCCCACCCTCCCAGTGC 118
        |||||||||||||||||||||||||||||| |||||||||||||||||||||||||
Db   5 CACACGCCGTCTTCTCTTTGCCTCAGCCACTTCCTTCCTTCCGCCTCACCTCCCACCCTCCCAGTGC 64

Qy 119 ACTGAAGAGAGGTAACCGGGTCCAGACCCACGCGCGCCAGTTCTCCGGCGGGAAGGAAAA 178
        |||||||||||||||||||||||||||||| |||||||||||||||||||||||||
Db  65 ACTGAAGAGAGGTAACCGGGTCCAGACCCACCGCGCGCCAGTTCTCCGGCGGGAAGGAAAA 124

Qy 179 CCGCGCAGAGAGGGCAGCAGCAATGAATGTGATCACGAGGTTAACCTCTTAGTGGAGGAAATT 238
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 125 CCGCGCAGAGAGGGCAGCAGCAATGAATGTGATCACGAGGTTAACCTCTTAGTGGAGGAAATT 184

Qy 239 CATCGTTTGGGTTCAAAAAATGCTGATGGAAAGTTAAGCGTGGAAATTTGGGGTCCTCTTC 298
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 185 CATCGTTTGGGTTCAAAAAATGCTGATGGAAAGTTAAGCGTGGAAATTTGGGGTCCTCTTC 244

Qy 299 CGTGATGATAAATGTGCCAACCTCTTGAAGCATTGGTAGGAACTCTTAAAGCTGCAAAA 358
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 245 CGTGATGATAAATGTGCCAACCTCTTGAAGCATTGGTAGGAACTCTTAAAGCTGCAAAA 304

Qy 359 CGAAGGAAGATTGTAACATATCCAGGAGAGCTGCTTCTGCAAGGTGTTCATGATGATGTT 418
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 305 CGAAGGAAGATTGTAACATATCCAGGAGAGCTGCTTCTGCAAGGTGTTCATGATGATGTT 364

Qy 419 GACATTATATATTACTGCAAGATTAATGTGGTTTACATATCTTTATGTACTGCCATTTTTG 478
        ||||||||| |||||||||||||||||||||||||||||||||||||||||||||||
Db 365 GACATTATATTACTGCAAGATTAATGTGGTTTACATATCTTTATGTACTGCCATTTTTTG 424

Qy 479 TTTCTGGTAAACTGGAATATAAAGTGAAAGAACAAACATTTGAACATACTTAATGTATTT 538
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 425 TTTCTGGTAAACTGGAATATAAAGTGAAAGAACAAACATTTGAACATACTTAATGTATTT 484

Qy 539 TTATAGAACTTGTAAACGAAAGGAGATTCATGTTTTAGAAGTCGTGTCCTTTTTTTATATC 598
```

SULT 9
521898
CUS   CD521898
FINITION AGENCOURT 14355142 NIH MGC 191 Homo sapiens cDNA clone
         IMAGE:30413687 5', mRNA sequence.
                                906 bp   mRNA   linear   EST 06-JUN-2003
CESSION CD521898
RSION   CD521898.1   GI:31453616
YWORDS   EST.
URCE   Homo sapiens (human)
ORGANISM Homo sapiens
         Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
         Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
FERENCE 1 (bases 1 to 906)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE   National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
MMENT   Contact: Daniela S. Gerhard, Ph.D.
         Office of Cancer Genomics

Query Match   80.4%;   Score 671.4;   DB 13;   Length 676;
Best Local Similarity   99.4%;   Pred. No. 1.6e-149;
Matches 672;   Conservative 0;   Mismatches 0;   Indels 0;   Gaps 0;

```
Qy    120 CTGAGAAGTAACCGGGTCCAGACCCACGGCGGCCCAGTCTCCGCGCGGGAAGGAAAAC 179
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 CTGAGAAGTAACCGGGTCCAGACCCACGGCGGCCCAGTCTCCGCGCGGGAAGGAAAAC 60

Qy    180 CGCGCAGAGAGGCAGCAATGAATGTGGATCACGAGGTTAACCTCTTAGTGGAGGAAATTC 239
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 CGCGCAGAGAGGCAGCAATGAATGTGGATCACGAGGTTAACCTCTTAGTGGAGGAAATTC 120

Qy    240 ATCGTTTGGGTTCAAAAAATGCTGATGGAAAGTTAAGCGTGGAAATTTGGGGTCCTCTTCC 299
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 ATCGTTTGGGTTCAAAAAATGCTGATGGAAAGTTAAGCGTGGAAATTTGGGGTCCTCTTCC 180

Qy    300 GTGATGATAAAATGTGCCAACCTCTTGAAGCATTGGTAGGAACTCTTAAAGCTGCAAAAC 359
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 GTGATGATAAAATGTGCCAACCTCTTGAAGCATTGGTAGGAACTCTTAAAGCTGCAAAAC 240

Qy    360 GAAGGAAGATTGTAACATATCCAGGAGAGAGCTGCTTCTGCGAGGTGTTCATGATGATGTTG 419
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 GAAGGAAGATTGTAACATATCCAGGAGAGAGCTGCTTCTGCGAAGGTGTTCATGATGATGTTG 300

Qy    420 ACATTATATTACTGCAAGATTAATGTGGTTTACATATCTTTATGTACTGCCATTTTTTGT 479
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 ACATTATATTACTGCAAGATTAATGTGGTTTACATATCTTTATGTACTGCCATTTTTTGT 360

Qy    480 TTCTGGTAAACTGGAATATAAAGTGAAAGAACAAACATTTGAACATACTTAATGTGATTTT 539
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 TTCTGGTAAACTGGAATATAAAGTGAAAGAACAAACATTTGAACATACTTAATGTATTTTT 420

Qy    540 TATAGAACTTTGTAAACGAAAGGAGATTCATGTTTAGAAGTCTGTCCTTTTTTATATCT 599
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 TATAGAACTTTGTAAACGAAAGGAGATTCATGTTTAGAAGTCTGTCCTTTTTTTATATCT 480

Qy    600 TGAAAGAAAATCTATGTATGATGCCTATAAAATAATCCTATTATTTTTCTCAGGAATCTG 659
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 TGAAAGAAAATCTATGTATGATGCCTATAAAATAATCCTATTATTTTTTCTCAGGAATCTG 540

Qy    660 GTTAGGAATTGCAGGCAATGAGATTTTTTGCGGGGATGGGAATGTTTGTTCATAA 719
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 GTTAGGAATTGCAGGCAATGAGATTTTTTGCGGGGATGGGAATGTTTGNTCATAA 600

Qy    720 ATAATTAGACATTTTCTATAGATATTTGACATTCTGCGAAAGCAACAAGCAAACTGAAGA 779
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 ATAATTAGACATTTTCTATAGATATTTGACATTCTGCGAAAGCAACAAGCAAACTGAAGA 660

Qy    780 CCAACTCCTATGAGAA 795
          ||||||||||||||||
Db    661 CCAACTNCTATGAGAA 676
```

RESULT 11
BU598760
LOCUS   BU598760   717 bp   mRNA   linear   EST 20-SEP-2002
DEFINITION AGENCOURT_8822409 NIH_MGC_142 Homo sapiens cDNA clone IMAGE:6457680
            5', mRNA sequence.
ACCESSION   BU598760
VERSION   BU598760.1   GI:23250519
KEYWORDS   EST.
SOURCE   Homo sapiens (human)
  ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 717)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE   National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT   Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: NCI

```
485 TTATAGAACTTTGTAAACGAAAGGAGATTCATGTTTAGAAGTCTGTCCTTTTTTATATC 544
    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

599 TTGAAAGAAAATCTATGTATGATGCCTATAAAATAATCCTATTATTTTTCTCAGGAATCT 658
    |||||||||||||||||||||||||||||||||||||||||||||||||||||||||

545 TTGAAAGAAAATCTATGTATGATGCCTATAAAATAATCCTATTATTTTTTCTCAGGAATCT 604

659 GGTTAGGAATTGCAGGCAATGAGATTTTTTGCGGGGATGGGAATGTTTGTTCATA 718
    ||||||||||||||||||||||||||||||||||||||||||||||||||||||

605 GGTTAGGAATTGCAGGCAATGAGATTTTTTGCGGGGATGGGAATGTTTGTTCATA 664

719 AATAATT-AGACATTTTCTATAGATATTTGA--CATTCTGCGAAAGCAACAACCACAAGCAAACTG 775
        |||||||| ||||||||||||||   |||||||||||||||||||||

665 AATAATTAAAACTTTTTCAATAAATATATTTGAACATTTCCGCGAAAGCAACCACAAGCAAACTG 724

776 AAGACCAACTCCTATGAGAAAATA 798
    ||||||||||| |||

725 GAGACCAACTCCCTAAGGAAAA 747
```

:SULT 10
:110370
LOCUS   BX110370   676 bp   mRNA   linear   EST 07-FEB-2003
DEFINITION BX110370 Soares_total_fetus_Nb2HF8_9w Homo sapiens cDNA clone
            IMAGP998B181944 ; IMAGE:788657, mRNA sequence.
ACCESSION   BX110370
VERSION   BX110370.1   GI:27878109
KEYWORDS   EST.
SOURCE   Homo sapiens (human)
  ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1   (bases 1 to 676)
  AUTHORS   Ebert,L., Heil,O., Hennig,S., Neubert,P., Partsch,E., Peters,M.,
            Radelof,U., Schneider,D. and Korn,B.
  TITLE   Human UnigeneSet - RZPD3
  JOURNAL   Unpublished (2003)
COMMENT   Contact: Ina Rolfs
            RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
            Im Neuenheimer Feld 580, D-69120 Heidelberg, Germany
            RZPD; IMAGp998B181944.
            RZPDLIB; I.M.A.G.E. cDNA Clone Collection;
            Human UnigeneSet - RZPD3 (RZPDLIB No.972)
            http://www.rzpd.de/CloneCards/cgi-
            bin/showLib.pl.cgi/response?libNo=972 Contact: Ina Rolfs
            RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
            Heubnerweg 6, D-14059 Berlin, Germany
            Tel: +49 30 32639 101
            Fax: +49 30 32639 111
            www.rzpd.de
            This clone is available royalty-free from RZPD;
            contact RZPD (clone@rzpd.de) for further information. Seq primer:
            M13r, Primer sequence: TTTCACAGGAAACAGCTATGAC.

FEATURES   Location/Qualifiers
     source   1..676
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGp998B181944 ; IMAGE:788657"
                     /dev_stage="8-9 weeks"
                     /lab_host="DH10B"
                     /clone_lib="Soares_total_fetus_Nb2HF8_9w"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
                     was prepared from mRNA obtained from pooled 8-9 week
                     (total) fetus material with a Not I - oligo(dT) primer [5'
                     TGTTACCAATCTGAAGTGGGAGCGGCGCCGCCTTAATTTTTTTTTTTTTTTTT 3'].
                     Double-stranded cDNA was ligated to Eco RI adaptors
                     (Pharmacia), digested with Not I and cloned into the Not I
                     and Eco RI sites of the modified pT7T3 vector. Library
                     went through one round of normalization, and was
                     constructed by Bento Soares and M. Fatima Bonaldo. "

ORIGIN
```

```
Db   481  TTCTCAGGAATCTGGTTAGGAATTGGAGGCAATGAGATTTTTGCGGGGGCAGGGATGGGA 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   706  ATGTTTGTTCATAAATAATTAGACATTTCTATAGATATTGACATTCTGCGAAAGCAAC 765
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   541  ATGTTTGTTCATAAATAATTAGACATTTCTATAGATATTGACATTCTGCGAAAGCAAC 600

Qy   766  AAGCAAACTGAAGACCAACTCCTATGAGAGAAATATTATGAGTTTATGTAATAAAGACATG 825
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   601  AAGCAAACTGAAGACCAACTCCTATGAGAGAAATATTATGAGTTTATGTAATAANAGACATG 660

Qy   826  TAACTGTCTT 835
              ||||||||||
Db   661  TAACTGTCTT 670
```

RESULT 12

```
LOCUS       BG530613                755 bp    mRNA     linear   EST 03-APR-2001
DEFINITION  602560008F1 NIH_MGC_61 Homo sapiens cDNA clone IMAGE:4698097 5',
            mRNA sequence.
ACCESSION   BG530613
VERSION     BG530613.1  GI:13522150
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 755)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
            cDNA Library Preparation: CLONETECH Laboratories, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Incyte Genomics, inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM1527  row: p  column: 02
            High quality sequence stop: 734.
FEATURES             Location/Qualifiers
     source          1..755
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4698097"
                     /tissue_type="embryonal carcinoma"
                     /clone_lib="NIH_MGC_61"
                     /note="Organ: testis; Vector: pDNR-LIB (Clontech); Site_1:
                     SfiI (ggccgcctcggcc); Site_2: SfiI (ggccattatggcc);
                     Double-stranded cDNA was prepared from cell line RNA.  5'
                     and 3' adaptors were used in cloning as follows: 5' adaptor
                     sequence: 5'-CACGGCCATTATGGCC-3' and 3' adaptor
                     sequence: 5'-ATTCTAGAGGCCGAGGCGGCCGACATG-dT(30)BN-3'
                     (where B = A, C, or G and N = A, C, G, or T). Average
                     insert size 1.75 kb (range 0.9-4.0 kb). 15/15 colonies
                     contained inserts by PCR. This library was enriched for
                     full-length clones and was constructed by Clontech
                     Laboratories (Palo Alto, CA). Note: this is a NIH_MGC
                     Library."
ORIGIN
```

Query Match   79.4%; Score 663; DB 12; Length 755;
Best Local Similarity 97.4%; Pred. No. 1.6e-147;
Matches 706; Conservative   0; Mismatches 15; Indels   4; Gaps   3;

```
Qy   60  ACAGCCGTCTTTCTTTGCCTCAGCCACTTCCTTCCTTCGCCTCACCCTCCCAGTGCA 119
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2  ACAGCCGTCTTTCTTTGCCTCAGCCACTTCCTTCCTTCGCCTCACCCTCCCAGTGCA 61
```

Second column:

cDNA Library Preparation: Michael Brownstein Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM2639 row: p column: 01
High quality sequence stop: 655.

```
FEATURES             Location/Qualifiers
     source          1..717
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6457680"
                     /tissue_type="mixed (pool of 40 RNAs)"
                     /lab_host="DH10B (T1-phage-resistant)"
                     /clone_lib="NIH_MGC_142"
                     /note="Vector: pDNR-LIB; Site_1: SfiI (ggccattatggcc);
                     Site_2: SfiI (ggccgcctcggcc); Double-stranded cDNA was
                     prepared from a pool of 40 cell line polyA+ RNAs (bladder
                     - 2%, blood - 33.4%, brain - 5.6%, breast - 12.5%, colon -
                     4%, connective tissue - 1.4%, eye - 1%, intestine - 2.6%,
                     kidney - 2.2%, liver - 5.7%, lung - 10.8%, NK-cell - 5.2%,
                     ovary - 4%, pharynx - 2.5%, prostate - 4.3%, salivary
                     gland - 1.3%, and skin - 2.3%). 5' and 3' adaptors were
                     used in cloning as follows:
                     5'-AAGCAGTGGTATCAACGCAGAGTGGCCATTACGGCCGGG-3' and
                     5'-ATTCTAGAGGCCGAGGCGGCCGACATG-dT(30)NN-3'. Full-length
                     enriched library was constructed using the Clontech
                     Creator SMART kit and size-selected to contain the >0.5 kb
                     size fraction (other fractions present in NIH_MGC 141).
                     Library created in the laboratory of M. Brownstein (NIMH,
                     NIH). Note: this is a NIH_MGC Library."
ORIGIN
```

Query Match   79.9%; Score 667.4; DB 13; Length 717;
Best Local Similarity 99.7%; Pred. No. 1.4e-148;
Matches 668; Conservative   0; Mismatches   2; Indels   0; Gaps   0;

```
Qy   166  GCGGGGAGGAAATTCATCGTTTGGGTTCAAAAAATGCTGATGGAAAGTTAAGCGTGAAATT 225
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1  GGGGGAGGAAATTCATCGTTTGGGTTCAAAAAATGCTGATGGAAAGTTAAGCGTGAAATT 60

Qy   226  AGTGGAGGAAATTCATCGTTTGGGTTCAAAAAATGCTGATGGAAAGTTAAGCGTGAAATT 285
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61  AGTGGAGGAAATTCATCGTTTGGGTTCAAAAAATGCTGATGGAAAGTTAAGCGTGAAATT 120

Qy   286  TGGGGTCCTCTTCCGTGATGATAAATGTGCCAACCTCTTTGAAGCATTGGTAGGAACTCT 345
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121  TGGGGTCCTCTTCCGTGATGATAAATGTGCCAACCTCTTTGAAGCATTGGTAGGAACTCT 180

Qy   346  TAAAGCTGCAAAACGAAGGAAGATTGTAACATATCCAGAGAGCTGCTTCTGCAAGGTGT 405
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181  TAAAGCTGCAAAACGAAGGAAGATTGTAACATATCCAGAGAGCTGCTTCTGCAAGGTGT 240

Qy   406  TCATGATGATGTTGACATTATTACTGCAAGATTAATGTGGTTTACATATCTTTATGTA 465
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241  TCATGATGATGTTGACATTATTACTGCAAGATTAATGTGGTTTACATATCTTTATGTA 300

Qy   466  CTGCCATTTTTGTTTCTGGTAAACTGAAACTGGAATTAAAAGTGAAAGAACAAACATTTGAACAT 525
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301  CTGCCATTTTTGTTTCTGGTAAACTGAAACTGGAATTAAAAGTGAAAGAACAAACATTTGAACAT 360

Qy   526  ACTTAATGTATTTTATAGAACTTGTAAACGAAAGGAAGAGATTCATGTTTTAGAAGTCTGT 585
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361  ACTTAATGTATTTTATAGAACTTGTAAACGAAAGGAAGAGATTCATGTTTTAGAAGTCTGT 420

Qy   586  CCTTTTTTATATCTTGAAAGAAAATCATGTATGATGCTATAAAATAAATCCTATTATTT 645
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421  CCTTTTTTATATCTTGAAAGAAAATCATGTATGATGCTATAAAATAAATCCTATTATTT 480

Qy   646  TTCTCAGGAATCTGGTTAGGAATTGGAGGCAATGAGATTTTTGCGGGGGCAGGGATGGGA 705
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
http://image.llnl.gov
Plate: LLCM3039  row:  1  column: 22
High quality sequence stop: 451.
                 Location/Qualifiers
FEATURES
     source    1..765
               /organism="Homo sapiens"
               /mol_type="mRNA"
               /db_xref="taxon:9606"
               /clone="IMAGE:6724487"
               /tissue_type="mixed (pool of 40 RNAs)"
               /lab_host="DH10B (T1-phage-resistant)"
               /clone_lib="NIH_MGC_126"
               /note="Vector: pDNR-LIB; Site_1: SfiI (ggccattatggcc);
               Site_2: SfiI (ggccgcctcggcc); Double-stranded cDNA was
               prepared from a pool of 40 cell line polyA+ RNAs (bladder
               - 2%, blood - 33.4%, brain - 5.6%, breast - 12.5%, colon -
               4%, connective tissue - 1.4%, eye - 1%, intestine - 2.6%,
               kidney - 2.2%, liver - 5.7%, lung - 10.8%, NK-cell -
               5.2%, ovary - 4%, pharynx - 2.5%, prostate - 4.3%,
               salivary gland - 1.3%, and skin - 2.3%).  5' and 3'
               adaptors were used in cloning as follows:
               5'-AAGCAGTGGTATCAACGCAGAGTGGCCATTACGGCCGGG-3' and
               5'-ATTCTAGAGGCGGCCGACATG-dT(30)NN-3'.  Full-length
               enriched library was constructed using the Clontech
               Creator SMART kit and size-selected to contain the 0.5-1
               kb size fraction (other fractions present in NIH_MGC_127
               and NIH_MGC_128)  Library created in the laboratory of T.
               Usdin, M.D., Ph.D.  (NIMH, NIH).  Note: this is a NIH_MGC
               Library."

ORIGIN

Query Match    78.8%;   Score 657.8;   DB 13;   Length 765;
Best Local Similarity   94.8%;   Pred. No. 2.8e-146;
Matches 680;  Conservative   0;  Mismatches 37;  Indels   0;  Gaps   0;

Qy  115 GTGCACTGAAGAAGGTAACCGGGTCCAGACCCACCGGGCGGCCAGTTCTCCGGCGCGGAAGG 174
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2 GGGCACTGAAGAAGGTAACCGGGTCCAGACCCACCGGGCGGCCAGTTCTCCGGCGCGGAAGG 61

Qy  175 AAAACCGCGAGAGGAGGCAGCAATGAATGTGGATCACGAGGTTAACCTCTTAGTGGAGGA 234
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   62 AAAACCGCGAGAGGAGGCAGCAATGAATGTGGATCACGAGGTTAACCTCTTAGTGGAGGA 121

Qy  235 AATTCATCGTTGGGTTCAAAAAATGCTGATGGAAAATGCTGAAAATTTGGGGTCCT 294
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  122 AATTCATCGTTGGGTTCAAAAAATGCTGATGGAAAATGCTGAAAATTTGGGGTCCT 181

Qy  295 CTTCCGTGATGATAAAATGTGCCAACCTCTTTGAAGCATTGGTAGGAACTCTTAAAGCTGC 354
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  182 CTTCCGTGATGATAAAATGTGCCAACCTCTTTGAAGCATTGGTAGGAACTCTTAAAGCTGC 241

Qy  355 AAAACGAAGGAAGATTGTAACATATCCAGGAGAGCTGCTTCTGCAAGGTGTTCATGATGA 414
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  242 AAAACGAAGGAAGATTGTAACATATCCAGGAGAGCTGCTTCTGCAAGGTGTTCATGATGA 301

Qy  415 TGTTGACATTATATTCTAAGATTAATGTGGTTTACATATCTTTAATGTACTGCCATTT 474
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  302 TGTTGACATTATATTCTAAGATTAATGTGGTTTACATATCTTTAATGTACTGCCATTT 361

Qy  475 TTTGTTTCTGGTAAACTGAAATAAATAAAGTGAAAGAACAAACATTTGAACATACTAATAGT 534
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  362 TTTGTTTCTGGTAAACTGAAATAAATAAAGTGAAAGAACAAACATTTGAACATACTAATAGT 421

Qy  535 ATTTTTATAGAACTTGTAAACGAAAGGAGATTCATGTTTAGAAGTCTGCGTCCTTTTTTA 594
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  422 ATTTTTATAGAACTTGTAAACGAAAGGAGATTCATGTTTAGAAGTCTGCGTCCTTTTTTA 481

Qy  595 TATCTTGAAAGAAAATCTATGTATGATGCTATAAAATAAAATCCTATTATTTTTCTCAGGA 654
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  482 TATCTTGAAAGAAAATCTATGTATGATGCTATAAAATAAAATCCTATTATTTTTCTCAGGA 541

Qy  655 ATCTGGTTAGGAATTGCAGGCAATGAGATTTTTTGCGGGCAGGGATGGAATGTTTGTT 714
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
120 CTGAAGAAGGTAACCGGGTCCAGACCCACCGGGCGGCCAGTTCTCCGGCGCGGAAGAAAC 179
    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 62 CTGAAGAAGGTAACCGGGTCCAGACCCACCGGGCGGCCAGTTCTCCGGCGCGGAAGAAAC 121

180 CGCGAGAGGCAGCAATGAATGTGGATCACGAGGTTAACCTCTTAGTGGAGGAAATTC 239
    |||||||||||||||||||||||||||||||||||||||||||||||||||||||
122 CGCGAGAGGCAGCAATGAATGTGGATCACGAGGTTAACCTCTTAGTGGAGGAAATTC 181

240 ATCGTTGGGTTCAAAAAATGCTGATGGAAAATTGGGGTCCTCCTCC 299
    |||||||||||||||||||||||||||||||||||||||||||||
182 ATCGTTGGGTTCAAAAAAATGCTGATGGAAAATTGGGGTCCTCCTCC 241

300 GTGATGATAAATGCCAACCTCTTTGAGCATGGTAGGAACTCTTAAAGCTGCAAAAC 359
    ||||||||||||||||||||||||||| ||||||||||||||||||||||||||
242 GTGATGATAAATGTGCCAACCTCTTTGAAGCATTGGTAGGAACTCTTAAAGCTGCAAAAC 301

360 GAAGGAGATTGTAACATATCCAGGAGAGCTGCTTCTGCAAGGTGTTCATGATGATGTG 419
    |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
302 GAAGGAGATTGTAACATATCCAGGAGAGCTGCTTCGCAAGGTGTTCATGATGATGTTG 361

420 ACATTATATTCTGAAGATTAATGTGGTTTACATATCTTTATGTACTGCCATTTTTTTGT 479
    ||||||||||| ||||||||||||||||||||||||||| |||||||||||  ||||||
362 ACATTATATTACTGCAAGATTAATGTGGTTTACATATCTTTATGTACTGCCA-TTTTTGT 420

480 TTCTGGTAAACTGGAATAAATAAAGTGAAAGAACAAACATTTGAACATACTAATGTATTTT 539
    |||||||||||||  |||  |||||||||||||||||||||||||||||||| ||||||||
421 TTCTGGTAAACTGGAATATACAGTGAAAGAACAAACATTTGAACATACTTAATGTATTTT 480

540 TATAGAACTTTGTAAACGAAAGGAGAGATTCATGTTTTAGAAGTCTGTCCTTTTTTATATCT 599
    ||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||
481 TATAGAACTTTGTAAACGAAAGGAGAGATTCATGTTTTAGAAGTCTGTCCTTTTTTTATATCT 540

600 TGAAAGAAAATCTATGTATGATGCTATAAAATAAATCCTATTATTTTTCTCAGGAATCTG 659
    ||||||  ||||||||||||||||||||||||||||||||||||||||||||||||||
541 TGAAAGACAATCTATGTATGATGCTATAAAATAAATCCTATTATTTTTCTCAGGAATCTG 600

660 GTTAGG-AATTGCAGGCAATGAGATTTTTTGCGGGCAGGGATGGGAATGTTTGTTCATA 718
    ||||||  |||||||||||||||||||    |||||| ||||||||||||| ||||||||
601 GTTAGGAAATTGCAGGCAATGAGATTT----GCGGGCATGGGAATGTTGTTCATA 660

719 AATAATTAGACATTTTCTATAGATATTTGACATTCTGCGAAAGCAACAAGCAAACTGAAG 778
    |||||||| |||||||||||||||||||| ||||||||  || |||||||||||||||
661 ACTAATTAGACCATTTCTATAGATATTTGACATTCT--GGAAGGCACAAGCAAACTGAAG 718

779 ACCAA 783
    |||||
719 ACCAA 723
```

RESULT 13
BU953438

LOCUS   BU953438   765 bp   mRNA   linear   EST 21-OCT-2002
DEFINITION AGENCOURT_10610027 NIH_MGC_126 Homo sapiens cDNA clone
            IMAGE:6724487 5', mRNA sequence.
ACCESSION   BU953438
VERSION   BU953438.1 GI:24183010
KEYWORDS   EST.
SOURCE   Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 765)
AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
TITLE   National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL   Unpublished (1999)
COMMENT   Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: NCI
            cDNA Library Preparation: Michael Brownstein Laboratory
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Agencourt Bioscience Corporation
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:

```
Qy  176 AAAACCGCGAGAGGAGGCAGCAATGAAATGTGGATCACGAGGTTAACCTCTTAGTGGAGGAA 235
            |||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db  122 AAAACCGCGAGAGGAGGCAGCAATGAAATGTGGATCACGAGGTTAACCCCTTAGTGGAGGAA 181

Qy  236 ATTCATCGTTTGGGTTCAAAAATGCTGATGGAAGTTAAGCGTGAAAATTGGGGTCCTC 295
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  182 ATTCATCGTTTGGGTTCAAAAATGCTGATGGAAGTTAAGCGTGAAAATTGGGGTCCTC 241

Qy  296 TTCCGTGATGATAAAATGTGCCAACCTCTTGAAGCATTGGTAGGAACTCTTAAAGCTGCA 355
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  242 TTCCGTGATGATAAAATGTGCCAACCTCTTGAAGCATTGGTAGGAACTCTTAAAGCTGCA 301

Qy  356 AAACGAAGGAAGATTGTAACATATCCAGGAGAGCTGCTTCTGCCAAGGTGTTCATGATGAT 415
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  302 AAACGAAGGAAGATTGTAACATATCCAGGAGAGCTGCTTCTGCCAAGGTGTTCATGATGAT 361

Qy  416 GTTGACATTATTACTGCAAGATTAATGTGGTTACATATCTTTATGTACTGCCATTT 475
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  362 GTTGACATTATTACTGCAAGATTAATGTGGTTACATATCTTTATGTACTGCCATTT 421

Qy  476 TTGTTTCTGGTAAACTGGAATATAAAGTGAAAGAACAAACATTGAACATACTAATGTA 535
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  422 TTGTTTCTGGTAAACTGGAATATAAAGTGAAAGAACAAACATTGAACATACTAATGTA 481

Qy  536 TTTTTATAGAACTTTGTAAACGAAAGGAGATTCATGTTTTAGAAGTCTGTCCTTTTTTAT 595
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  482 TTTTTATAGAACTTTGTAAACGAAAGGAGATTCATGTTTTAGAAGTCTGTCCTTTTTTAT 541

Qy  596 ATCTTGAAAGAAAATCTATGTATGATGCTATAAAATAAAATCCTATTATTTTTCTCAGGAA 655
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  542 ATCTTGAAAGAAAATCTATGTATGATGCTATAAAATAAAATCCTATTATTTTTCTCAGGAA 601

Qy  656 TCTGGT-AGGAATTGCAGGCAATGAGATTTTTGCGGGCCAGGGATGGGAA--TGTTTG 712
            ||||||  |||||||||||||||||||||||||||||||||||||||||||  |||||
Db  602 TCTGGTTAAGGAATTGCAGGCAATGAGATTTTTGCGGGCCAGGGATGGGAAATGGTTTG 661

Qy  713 TTCATAAATA-ATTAGACATTTTCTATAGA-TATTGACATTCT-GGCGAAAGCAACAAGC 769
            |||||  | |  |||||||||||||||||||  ||||||||||  ||||||||||||||||
Db  662 TTCATTAACATATTAGACATTTTCTATAGACCTTTTGACATTTCTGGCGAAAGCAACAAGC 721

Qy  770 AAACTGAAGACCAA 783
            ||||||||||||||
Db  722 AAACTGAAGACCA 735
```

RESULT 15
BG744405
LOCUS   BG744405   684 bp   mRNA   linear   EST 15-MAY-2001
DEFINITION 602723144F1 NIH_MGC_106 Homo sapiens cDNA clone IMAGE:4849571 5', mRNA sequence.
ACCESSION   BG744405
VERSION   BG744405.1   GI:14055058
KEYWORDS   EST.
SOURCE   Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 684)
AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
TITLE   National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL   Unpublished (1999)
COMMENT   Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Dr. Daniel McVicar, DBS/NCI
            cDNA Library Preparation: Ling Hong/Rubin Laboratory
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM1689 row: g column: 12
            High quality sequence stop: 682.
FEATURES   Location/Qualifiers

```
Qy  542 ATCTGGTTAGGAATTGCAGGCAATGAGATTTTTGCGGGCCAGGGATGGGAATGTTTGTT 601
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  715 CATAAATAATTAGACATTTTCTATAGATATTTGACATTCTGCGAAAGCAACAAACT 774
Db  602 CATAAATAATTAGACATTTTCTATAGATATTTGACATTCTGCGAAAACAACAAGCCAACC 661

Qy  775 GAAGACCAACTCCCTATGCAGGAGAATTATTATGCGGAGATGGTTAATGTAATAAGGACACATG 831
            |||||||| |||||||||||||||||||||||||||  ||||||||||||||||||||||||
Db  662 TGAAGACCACTCCCTATGCAGGAGAATTATTATGCGGAGATTATTATGCGGTGTAATAAGGACACATG 718
```

RESULT 14
BG502291
LOCUS   BG502291   747 bp   mRNA   linear   EST 27-MAR-2001
DEFINITION 602550569F1 NIH_MGC_61 Homo sapiens cDNA clone IMAGE:4658032 5', mRNA sequence.
ACCESSION   BG502291
VERSION   BG502291.1   GI:13463808
KEYWORDS   EST.
SOURCE   Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 747)
AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
TITLE   National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL   Unpublished (1999)
COMMENT   Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
            cDNA Library Preparation: CLONETECH Laboratories, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM1451 row: j column: 17
            High quality sequence stop: 742.
FEATURES   Location/Qualifiers
     source 1..747
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone="IMAGE:4658032"
            /tissue_type="embryonal carcinoma"
            /lab_host="DH10B (T1 phage-resistant)"
            /clone_lib="NIH_MGC_61"
            /note="Organ: testis; Vector: pDNR-LIB (Clontech); Site_1:
            SfiI (ggccgctcggcc); Site_2: SfiI (ggccattatggcc);
            Double-stranded cDNA was prepared from cell line RNA. 5'
            and 3' adaptors were used in cloning as follows: 5'
            adaptor sequence: 5'-CACGGCCATTATGGCC-3' and 3' adaptor
            sequence: 5'-ATTCTAGAGGCCGAGGCGGCCGACATG-dT(30)BN-3' Average
            (where B = A, C, or G and N = A, C, G, or T). Average
            insert size 1.75 kb (range 0.9-4.0 kb). 15/15 colonies
            contained inserts by PCR. This library was enriched for
            full-length clones and was constructed by Clontech
            Laboratories (Palo Alto, CA). Note: this is a NIH_MGC
            Library."

ORIGIN

Query Match   78.8%;   Score 657.6;   DB 12;   Length 747;
Best Local Similarity 98.0%;   Pred. No. 3.1e-146;
Matches 719;   Conservative 0;   Mismatches 9;   Indels 9;   Gaps 5;

```
Qy   56 CTCCACAGCCGTCTTTCTCTTTGCCCTCAGCCACTTCCTTCCTTCCGGCTCACCCTCCCAG 115
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2 CTCCACAGCCGTCTTTCTCTTTGCCCTCAGCCACTTCCTTCCTTCCGGCTCACCCTCCCAG 61

Qy  116 TGCACTGAAGAAGGTAACCGGGTCCAGACCCACGCGGCGGCCGCCCAGTTCTCCGGCGGAAGGA 175
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   62 TGCACTGAAGAAGGTAACCGGGTCCAGACCCACGCGGCGGCCGCCCAGTTCTCCGGCGGAAGGA 121
```

source   1..684
                /organism="Homo sapiens"
                /mol_type="mRNA"
                /db_xref="taxon:9606"
                /clone="IMAGE:4849571"
                /tissue_type="natural killer cells, cell line"
                /lab_host="DH10B (phage-resistant)"
                /clone_lib="NIH_MGC_106"
                /note="Organ: blood; Vector: pOTB7; Site_1: XhoI; Site_2:
                EcoRI; cDNA made by oligo-dT priming. Directionally cloned
                into EcoRI/XhoI sites using the following 5' adaptor:
                GGCACGAG(G). Library constructed by Ling Hong in the
                laboratory of Gerald M. Rubin (University of California,
                Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
                Superscript II RT (Life Technologies). Note: this is a
                NIH_MGC Library."

ORIGIN

Query Match   78.7%;   Score 657;   DB 12;   Length 684;
Best Local Similarity   99.3%;   Pred. No. 4.3e-146;
Matches   682;   Conservative   0;   Mismatches   0;   Indels   5;   Gaps   2;

        120 CTGAAGAAGGTAACCGGGTTCCAGACCCAGTTCCCGGCGGGAAGGAAAAC 179
             ||||||||||||||||||||||||||||||||||||||||||||||||||
          2 CTGAAGAAGGTAACCGGGTTCCAGACCCAGTTCCCGGCGGGAAGGAAAAC 61

        180 CGCCCAGAGAGGCAGCAATGAATGTGGATCACGAGGTTAACCTCTTAGTGGAGGAAATTC 239
             |||||||||| ||||||||||||||||||||||||||||||||||||||||
         62 CGCCCAGAG----AGCAATGAATGTGGATCACGAGGTTAACCTCTTAGTGGAGGAAATTC 117

        240 ATCGTTTGGGTTCAAAAAATGCTGATGAAAGTTAAGCGTGAAATTTGGGGTCCTCTTCC 299
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
        118 ATCGTTTGGGTTCAAAAAATGCTGATGAAAGTTAAGCGTGAAATTTGGGGTCCTCTTCC 177

        300 GTGATGATAAATGTGCCAACCTCTTGAAGCATTGGTAGGAACTCTTAAAGCTGCAAAAC 359
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
        178 GTGATGATAAATGTGCCAACCTCTTGAAGCATTGGTAGGAACTCTTAAAGCTGCAAAAC 237

        360 GAAGGAAGATTGTAACATATCCAGGAGAGCTGCTTCTCCAAGGTGTTCATGATGATGTTG 419
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
        238 GAAGGAAGATTGTAACATATCCAGGAGAGCTGCTTCTCGCAAGGTGTTCATGATGATGTTG 297

        420 ACATTATATTACTGCAAGATTAATGTGGTTTACATATCTTATGTACTGCCATTTTTTGT 479
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
        298 ACATTATATTACTGCAAGATTAATGTGGTTTACATATCTTATGTACTGCCATTTTTTGT 357

        480 TTCTGGTAAACTGGAATATAAAGTGAAAGAACAAACATTGAACATACTTAATGTATTTT 539
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
        358 TTCTGGTAAACTGGAATATAAAGTGAAAGAACAAACATTGAACATACTTAATGTATTTT 417

        540 TATAGAACTTTGTAAACGAAAGGAGAGATTCATGTTTAGAAGTCTGTCCTTTTTATATCT 599
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
        418 TATAGAACTTTGTAAACGAAAGGAGAGATTCATGTTTAGAAGTCTGTCCTTTTTATATCT 477

        600 TGAAAGAAAATCTATGTATGATGCTATAAAATAAATAATCCTATTATTTTTCTCAGGAATCTG 659
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
        478 TGAAAGAAAATCTATGTATGATGCTATAAAATAAATAATCCTATTATTTTTCTCAGGAATCTG 537

        660 GTTAGGAATTGCAGGCAATGAGATTTTTGCGGGGCAGGGATGGGAAATGTTTGTTCAT-A 718
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
        538 GTTAGGAATTGCAGGCAATGAGATTTTTGCGGGGCAGGGATGGGAAATGTTTGTTCAT AA 597

        719 AATAATTAGACATTTTCTATAGATATTTGACATTCGCGAAAGCAACAAGCAAACTGAAG 778
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
        598 AATAATTAGACATTTTCTATAGATATTTGACATTCGCGAAAGCAACAAGCAAACTGAAG 657

        779 ACCAACTCCTATGAGAAAATATTATGAT 805
             ||||||||||||||||||||||||||
        658 ACCAACTCCTATGAGAAATATTATGAT 684

Search completed: March 17, 2004, 14:23:43
Job time : 2826 secs

```
                    GenCore version 5.1.6
          Copyright (c) 1993 - 2004 Compugen Ltd.

nucleic - nucleic search, using sw model

on:    March 17, 2004, 12:24:02 ; Search time 91 Seconds
                         (without alignments)
                    5092.132 Million cell updates/sec

le:    US-09-648-310-3
fect score: 835
uence: 1 ggcacgaggcttgagcggcag.........taaagacatgtaactgctctt 835

ring table: IDENTITY_NUC
       Gapop 10.0 , Gapext 1.0

rched:    682709 seqs, 27747446 residues

al number of hits satisfying chosen parameters:    1365418

imum DB seq length: 0
imum DB seq length: 2000000000

st-processing: Minimum Match 0%
               Maximum Match 100%
               Listing first 45 summaries

abase ;  Issued_Patents NA:*
    1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
    2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
    3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
    4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
    5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
    6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.
```

SUMMARIES

sult No.	Score	% Query Match	Length	DB	ID	Description
1	747	89.5	786	4	US-09-023-655-79	Sequence 79, Appl
2	310	37.1	800	4	US-09-644-460-26	Sequence 26, Appl
3	186	22.3	219	4	US-09-621-976-17956	Sequence 17956, A
4	170.2	20.4	611	3	US-09-385-982-376	Sequence 376, App
5	50	6.0	6020	4	US-10-204-708-8	Sequence 8, Appli
6	45.2	5.4	177	2	US-09-032-684-20	Sequence 20, Appl
7	45.2	5.4	177	4	US-09-644-460-20	Sequence 20, Appl
8	45.2	5.4	832	4	US-09-621-976-2813	Sequence 2813, Ap
9	44.8	5.4	640681	4	US-09-790-988-1	Sequence 1, Appli
10	44.4	5.3	1218	4	US-08-731-722-4	Sequence 4, Appl
11	43.8	5.2	8537	4	US-10-204-708-42	Sequence 42, Appl
12	43	5.1	1189	4	US-09-221-017B-77	Sequence 77, Appl
13	42.8	5.1	2861	1	US-08-299-953-1	Sequence 1, Appli
14	42.8	5.1	2861	1	US-08-459-415-1	Sequence 1, Appli
15	42.8	5.1	2861	1	US-09-066-687-1	Sequence 1, Appli
16	42.8	5.1	3881	5	PCT-US95-11231-1	Sequence 1, Appli
17	42.8	5.1	3881	1	US-08-299-953-2	Sequence 2, Appli
18	42.8	5.1	3881	4	US-08-459-415-2	Sequence 2, Appli
19	42.8	5.1	3881	4	US-09-066-687-2	Sequence 2, Appli
20	42.8	5.1	3881	5	PCT-US95-11231-2	Sequence 2, Appli
21	42.4	5.1	640681	4	US-09-790-988-1	Sequence 1, Appli
22	41.4	5.0	487	4	US-09-621-976-16334	Sequence 16334, A
23	41.4	5.0	1497	4	US-09-220-132-94	Sequence 94, Appl
24	41	4.9	6156	4	US-09-066-687-59	Sequence 59, Appl
25	40.4	4.8	6113	4	US-10-204-708-13	Sequence 13, Appl
26	40.2	4.8	3095	6	5231168-1	Patent No. 5231168
27	40.2	4.8	7101	1	US-08-480-604A-9	Sequence 9, Appli
28	40.2	4.8	7101	2	US-08-405-496A-9	Sequence 9, Appli
29	40.2	4.8	7101	3	US-08-915-136-9	Sequence 9, Appli
30	40.2	4.8	7101	4	US-08-957-310-9	Sequence 9, Appli
31	40.2	4.8	7101	4	US-10-011-366-9	Sequence 9, Appli
32	40.2	4.8	7101	4	US-09-084-517-9	Sequence 9, Appli
33	40	4.8	4702	4	US-08-956-171E-268	Sequence 268, App
34	40	4.8	7218	1	US-08-232-463-14	Sequence 14, Appl
35	39.8	4.8	1182	4	US-09-134-000C-3346	Sequence 3346, Ap
36	39.8	4.7	501	c	US-08-916-421B-1	Sequence 1, Appli
37	39.6	4.7	6306	4	US-08-601-198-81	Sequence 81, Appl
38	39.6	4.7	1110	4	US-10-204-708-50	Sequence 50, Appl
39	39.4	4.7	8607	c	US-09-328-352-2373	Sequence 2373, Ap
40	38.8	4.6	9636	1	US-10-204-708-71	Sequence 71, Appl
41	38.8	4.6	9636	4	US-08-323-170B-1	Sequence 1, Appli
42	38.8	4.6	9636	4	US-08-954-441-1	Sequence 1, Appli
43	38.8	4.6	11049	4	US-10-204-708-22	Sequence 22, Appl
44	38.6	4.6	11049	4	US-10-204-708-23	Sequence 23, Appl
45	38.2	4.6	3001	4	US-09-539-33D-146	Sequence 146, App

ALIGNMENTS

```
RESULT 1
US-09-023-655-79
; Sequence 79, Application US/09023655
; Patent No. 6607879
; GENERAL INFORMATION:
;  APPLICANT: Cocke, Benjamin G.
;  APPLICANT: Susan G. Stuart
;  TITLE OF INVENTION: Jeffrey J. Seilhamer
;  TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE
;  TITLE OF INVENTION: EXPRESSION
;  NUMBER OF SEQUENCES: 1508
;  CORRESPONDENCE ADDRESS:
;   ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
;   STREET: 3174 PORTER DRIVE
;   CITY: PALO ALTO
;   STATE: CALIFORNIA
;   COUNTRY: USA
;   ZIP: 94304
; COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/09/023,655
;   FILING DATE: HEREWITH
;   CLASSIFICATION:
; PRIOR APPLICATION DATA:
;   APPLICATION NUMBER:
;   FILING DATE:
;   CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
;   NAME: Zeller, Karen J.
;   REGISTRATION NUMBER: 37,071
;   REFERENCE/DOCKET NUMBER: PA-0001 US
; TELECOMMUNICATION INFORMATION:
;   TELEPHONE: (650) 855-0555
;   TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 79:
;  SEQUENCE CHARACTERISTICS:
;   LENGTH: 786 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: single
;   TOPOLOGY: linear
;  IMMEDIATE SOURCE:
;   LIBRARY: THP1PLB01
;   CLONE: 012364
US-09-023-655-79

  Query Match             89.5%;  Score 747;  DB 4;  Length 786;
```

```
;  PRIOR APPLICATION NUMBER: US 09/197,889
;  PRIOR FILING DATE: 1998-11-23
;  PRIOR APPLICATION NUMBER: US 09/185,115
;  PRIOR FILING DATE: 1998-11-03
;  PRIOR APPLICATION NUMBER: US 09/032,684
;  PRIOR FILING DATE: 1998-02-27
;  NUMBER OF SEQ ID NOS: 42
;  SOFTWARE: FastSEQ for Windows Version 4.0
SEQ ID NO 26
;  LENGTH: 800
;  TYPE: DNA
;  ORGANISM: homo sapiens
US-09-644-460-26

Query Match            37.1%;    Score 310;  DB 4;  Length 800;
Best Local Similarity  68.7%;    Pred. No. 7.1e-72;
Matches  537;  Conservative   0;  Mismatches  220;  Indels  25;  Gaps   7;

Qy    43 CCTGCTCCTCCTCCCAGCGCCGTCTTTCTCCTTGCCTCAGCCACTTCCTTCCTTCGCC 102
         |||   ||||||| | | | || || |||||||||  ||||| |  |||||||||||| ||
Db    16 CGTCCCCTCCCTTCCTCCACTGCAGCCTTTCTCTTTAGCCCGAACCACTTCCTTCCTTCTGCT 75

Qy   103 TCACCCTCCCAGTGCACTGAGAAGTAACCGGGTCAGACGCGGGCGAGCAGCTC 162
         |||     |||   ||| | ||||||||| |||| |  ||||| || |||||
Db    76 TGTTCCTCCCTAGGGCGGCGGAAGTGAGTGCAGGGTCAGACCCACGGGCGAGCAGCTC 135

Qy   163 CCGGGCGGGAGGGAAAACCCGCGCAGAGAGGCAGCAATGAATGTGGATCACGAGGTTAACCT 222
         | |||  | ||||| |||  || |||||| |||||||||||| || ||||| ||||||||| |
Db   136 TTCAGTGAAGAAGAAGCAATCGGAGGGTCAGCAATGAACGTGGAGCATGGAAAGTTAACCT 195

Qy   223 CTTAGTGGAGGAAATTCATCGTTGGGTCAAAAAAATGCTGATGGAAAGTAAGCGTGAA 282
         |  |||||| |||||||||||| |||||| || ||||| ||||| ||| |||| |||||
Db   196 CCTGGTGGAGGAAATTCATCGTCTGGGTTCCAAAAATGCCGATGGGAAACTGAGTGTGAA 255

Qy   283 ATTTGGGGTCCTCTCCGTGATGATAAATGTGCAACCTCTTG-AAGCATTGGTAGGAA 341
         ||| ||||| ||||||| |||| | ||||| ||||||||||  |||| ||| |||||
Db   256 GTTTGGGGTCCTCTTCCAGACGCACAGATGTGCCAATCTCTTTGAAACCGTTGGTGGGAA 315

Qy   342 CTCTTAAAG-CTGCAAAACGAGGAAGATTGTAACATATCCAGGAGAGCTGCTTCTGCAA 400
         ||||| |||  ||||| ||| |||| ||||| || || |||||||||| || || |||
Db   316 CTCTGAAAGCCCGCAAAACGAGGAAGAAGATTGTTACGTACGTAACGATCTGG 375

Qy   401 GGTGTTCATGATGATGTGACATTATATTACTGCAAGATTAATGTGGTTTACATATCTTT 460
         || |||||||| ||||| |||||||| || |||||| || | |||||| |||| |||||
Db   376 GGTGTTCATGATGATGTTGACATTGTATTGTGCAAGATTAATGTGTTTTGCAGATCTGG 435

Qy   461 ATGTACTGCCATTTTTGTTTCTGTAAACTGGAAT-ATAAGTGAAAGAACAAACATTT 519
         | ||   ||   |||  | |||| |||| |||||   |||| |||   | ||||||||
Db   436 GGGTA------------TCTGGTAAAACTGGAATAAATTAAGTTAAGGACAAACAT-- 478

Qy   520 GAACATACTTAATGTATTTTATAGAACTTTGTAAACGAAAGGAGATTCATGTTTAGAA 579
         ||| || |  || |||  |||||| | |||| || ||||  |||| | ||| |||||
Db   479 -GAAGTTCCTTATGTATTTTTATAGACCCTTGTAAACAAAGGGGA--CTTGTTGAGAAG 535

Qy   580 GTCTGTCCTTTTTTATATCTTGAAGAAAATCATGTGATGCTATAAAATAAATCCTA 639
         || |||| |||| || |||| || |  |   || | || |||||| |||  |||||
Db   536 TCCTGTTTTTATACCTTGGAGCAAAACATTACAATGTAAAAATAAACAAAACCTGTTATT 595

Qy   640 TTATTTTCTCAGAAATCTGGTTAGGAATCGGGAGACGTAGGCAATAAAATGTTTTCAGAGGTGCGA 699
         | |||||  || ||||   ||  | | || | | || |||| ||||| ||| | || | |||| ||
Db   596 TTTTTTTTCTTAAGAAGTTAATCGGGAGACGATTAGATATTGGAGCAATAAAATGTTTTCAGAGGTGCGA 655

Qy   700 ATGGGAAATGTTGTCATAAATAATTAGACATTTCTATAGATATTTGACATTCTCGCGAA 759
         ||| ||| | | | ||||| ||| || | ||| | || || | || ||  || | |||
Db   656 AAAAGTTTTGTTTTTTAAACCATTCTTAGTCT--CTGCCACACTTCCTCTGTAATAAA 713

Qy   760 AGCAACAAGCAAACTGAAGACCAACTCCTATGAGAAAATCATTATGATGTTTATGTAATAAA 819
         | ||| ||||| | | |||| |||| |||| |||| || ||| | |||||| |  |||||||
Db   714 AGTGAGAAGGCAACTAAAGACCAACTCCTGGGTGGAAAAATTATTATGTTTTATGTAATAAAAAAA 773

Qy   820 GA 821
         ||
Db   774 AA 775
```

```
Best Local Similarity   99.0%;   Pred. No. 2.1e-186;
Matches  773;  Conservative   0;  Mismatches   0;  Indels  3;  Gaps  2;

Qy    58 CCACAGCCGTCTTTCTCCTTGCCTCAGCCACTTCCTTCCTTCGCCTCACCCTCCCCAGTG 117
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1 CCACAGCCGTCTTTCTCCTTGCCTCAGCCACTTCCTTCCTTCGCCTCACCCTCCCCAGTG 60

Qy   118 CACTGAGAAGTAACCGGGTCCAGACCCACGGCGCGCCCAGTTCTTCCGCGCGGGAAGGAAA 177
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61 CACTGAGAAGTAACCGGGTCCAGACCCACGGCGCGCCCAGTTCTTCCGCGCGGGAAGGAAA 120

Qy   178 ACCGGCGCAGAGAGGCAGCAATGAATGTGGATCACGAGGTTAACCTCTTAGTGAGGAAAT 237
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 ACCGGCGCAGAGAGGCAGCAATGAATGTGGATCACGAGGTTAACCTCTTAGTGAGGAAAT 180

Qy   238 TCATCGTTGGGTTCAAAAAAATGCTGATGGAAAGTAAGC--GTGAAATTTGGGGTCCTC 295
         |||||||||||||||||||||||||||||||||||||||  |||||||||||||||||||
Db   181 TCATCGTTGGGTTCAAAAAAATGCTGATGGAAAGTAAGCTTTTGGGGTCCTC 240

Qy   296 TCCGTGATGATAAATGTGCCAACCTCTTGAAGCATTGGTAGGAACTCTTAAAGCTGC- 354
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 TTCCGTGATGATAAATGTGCCAACCTCTTGAAGCATTGGTAGGAACTCTTAAAGCTGCA 300

Qy   355 AAAACGAAGGAAGATTGTAACATATCCAGGAGAGCTGCTTCTGCAAGATGTGTTCATGATGA 414
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 AAAACGAAGGAAGATTGTAACATATCCAGGAGAGCTGCTTCTGCAAGGTGTGTTCATGATGA 360

Qy   415 TGTTGACATTATTACTGCAAGATTAATGTTTACATATCTTTATGTACTGCCCATTT 474
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 TGTTGACATTATTACTGCAAGATTAATGTGGTTTACATATCTTTATGTACTGCCCATTT 420

Qy   475 TTTGTTTCTGGTAAACTGGAATGAATATAAGTGAAAGAACAAACATTGAACATACTTAATGT 534
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 TTTGTTTCTGGTAAACTGGAATGAATATAAGTGAAAGAACAAACATTGAACATACTTAATGT 480

Qy   535 ATTTTTATAGAACTTTGTAAACGAAAGGAGATTCATGTTTAGAAGTCTGTCGTCTCCTTTTTA 594
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481 ATTTTTATAGAACTTTGTAAACGAAAGGAGATTCATGTTTAGAAGTCTGTCGTCTCCTTTTTA 540

Qy   595 TATCTTGAAAGAAAATCTATGTGTATGATGCTATAAAAATAAATCCTATTATTTTTCTCAGGA 654
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   541 TATCTTGAAAGAAAATCTATGTGTATGATGCTATAAAAATAAATCCTATTATTTTTCTCAGGA 600

Qy   655 ATCTGGTTAGGAATTGCAGGCAATGAGAGATTTTTGCGGGCCAGGGAATGTTTGTT 714
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   601 ATCTGGTTAGGAATTGCAGGCAATGAGAGATTTTTGCGGGCCAGGGAATGTTTGTT 660

Qy   715 CATAAATAATTAGACATTTCTATAGATATTTGACATTCTCGCGAAAGCAACAAGCAAACT 774
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   661 CATAAATAATTAGACATTTCTATAGATATTTGACATTCTCGCGAAAGCAACAAGCAAACT 720

Qy   775 GAAGACCAACTCCTATGAGAAAATATTATGATGTTTATGTAATAAAGATGTAACTGTCT 834
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   721 GAAGACCAACTCCTATGAGAAAATATTATGATGTTTATGTAATAAAGATGTAACTGTCT 780

Qy   835 T 835
         |
Db   781 T 781
```

ESULT 2
S-09-644-460-26, Application US/09644460
Sequence 26, Application US/09644460
Patent No. 6657053
GENERAL INFORMATION:
APPLICANT: Fisher, Paul B.
TITLE OF INVENTION: Reciprocal Subtraction Differential
TITLE OF INVENTION: Display
FILE REFERENCE: 34587-C-PCT-USA
CURRENT APPLICATION NUMBER: US/09/644,460
CURRENT FILING DATE: 2000-08-23
PRIOR APPLICATION NUMBER: PCT/US99/04323
PRIOR FILING DATE: 1999-02-26

ULT 3
09-621-976-17956
equence 17956, Application US/09621976
atent No. 6639063
ENERAL INFORMATION:
APPLICANT: Dumas Milne Edwards, J.B.
APPLICANT: Jobert, S.
APPLICANT: Giordano, J.Y.
TITLE OF INVENTION: ESTs and Encoded Human Proteins.
FILE REFERENCE: GENSET.054PR2
CURRENT APPLICATION NUMBER: US/09/621,976
CURRENT FILING DATE: 2000-07-21
NUMBER OF SEQ ID NOS: 19335
SOFTWARE: Patent.pm
EQ ID NO 17956
LENGTH: 219
TYPE: DNA
ORGANISM: Homo sapiens
09-621-976-17956

uery Match   22.3%;   Score 186; DB 4;   Length 219;
est Local Similarity   99.5%;   Pred. No. 1.3e-39;
atches   197;   Conservative   0;   Mismatches   0;   Indels   1;   Gaps   1;

637 CTATTATTTTTCTCAGGAATCTGGTTAGGAATTGCAGGCAATGAGATTTTTGCGGGGCA 696
     |||||||||||||| |||||||||||||||||||||||||||||||||||||||||||||
1 CTATTATTTTTCTCAGGAATCTGGTTAGGAATTGCAGGCAATGAGATTTTTGCGGGGCA 60

697 GGGATGGGAATGTTGTTCATAAATAATTAATTAGACATTTCTCATAGATATTTGACATTCTGC 756
     ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
61 GGGATGGGAATGTTGTTCATAAATAATTAGACATTTCTCATAGATATTTGACATTCTGC 120

757 GAAAGCAACAAGCAAACTGAAGACCAACTCCTATGAGAAATATTATGATGTTTATGTAAT 816
     |||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
121 GAAA-CAACAAGCAAACTGAAGACCAACTCCTATGAGAAATATTATGATGTTTATGTAAT 179

817 AAAGACATGTAACTGTCT 834
     ||||||||||||||||||
180 AAAGACATGTAACTGTCT 197

ULT 4
-09-385-982-376
Sequence 376, Application US/09385982
Patent No. 6262334
GENERAL INFORMATION:
APPLICANT: ENDEGE, WILSON O., ET AL.
TITLE OF INVENTION: NOVEL HUMAN GENES AND GENE EXPRESSION
TITLE OF INVENTION: PRODUCTS: II
FILE REFERENCE: CCDNA-260XX
CURRENT APPLICATION NUMBER: US/09/385,982
CURRENT FILING DATE: 1999-08-30
EARLIER APPLICATION NUMBER: 09/328,111
EARLIER FILING DATE: 1999-06-08
EARLIER APPLICATION NUMBER: 60/117,393
EARLIER FILING DATE: 1999-01-27
EARLIER APPLICATION NUMBER: 60/098,639
EARLIER FILING DATE: 1998-08-31
NUMBER OF SEQ ID NOS: 544
SOFTWARE: FastSEQ for Windows Version 3.0
SEQ ID NO 376
LENGTH: 611
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)...(611)
OTHER INFORMATION: n = A,T,C or G
-09-385-982-376

Query Match   20.4%;   Score 170.2;   DB 3;   Length 611;

Best Local Similarity   78.2%;   Pred. No. 2.8e-35;
Matches   204;   Conservative   0;   Mismatches   56;   Indels   1;   Gaps   1;

Qy 461 ATGTACTGCCATTTTTTTGTTTCTGGTAAACTGGAATATAAAGTGAAAGAACAAACATTTG 520
      |||| | || ||||| ||||| |||||||||| ||| || || |||||| ||| |||||||
Db 308 ANGTNCNGCCATTTTTNGTTTCNGGTAAACNGGAATATANGNGAAAGAACAAACNTTGG 367

Qy 521 AACATACTTAATGTATTTTTATAGAACTTTGTAAACGAAGGAGATTCATGTTTTAGAAG 580
      |||||| || ||||||||||||||||||| || |||||| |||| || ||||||||||||
Db 368 AACATACTTAATGGATTTTTATAGAACTTTGNAAACCAAAGGAGATTCATGTTTANAAG 427

Qy 581 TCTGTCCTTTTTTATATCTTGAAAGAAAATCTATGTATGCTATAAAATAAATCCTAT 640
      |||||| |||||||||||||||||||||||| |||||||||||||||||||||||||
Db 428 TCTGGCCTTTTTATATCTTGGAAGAAAAT-TATGTNTGGAGGCTNTAAATAAATCCCAT 486

Qy 641 TATTTTTCTCAGGAATCTGGTTAGGAATTGCAGGCAATGAGATTTTTGCGGGCAGGGA 700
      ||||||||||||||||||| ||||||||||||||||||||||||| ||||||||||||
Db 487 TATTTTTCTCAGGAATCTGGTAGGGAGGTAGGAATTGCCGGCATGGGAANTTTTNNGGGGCCGGGATN 546

Qy 701 TGGGAATGTTTGTTCATAAAT 721
      |||||||||||| |||| |||
Db 547 GGAAAGTTTGGCCTAANAAAT 567

RESULT 5
US-10-204-708-8
; Sequence 8, Application US/10204708
; Patent No. 6677731
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA Replication
; TITLE OF INVENTION: by Assessing DNA Methylation
; FILE REFERENCE: 5013.1012
; CURRENT APPLICATION NUMBER: US/10/204,708
; CURRENT FILING DATE: 2003-05-06
; PRIOR APPLICATION NUMBER: PCT/EP01/03971
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: DE 10019058.8
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: DE 100019173.8
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 98
; SEQ ID NO 8
; LENGTH: 6020
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-204-708-8

Query Match   6.0%;   Score 50; DB 4; Length 6020;
Best Local Similarity   51.1%;   Pred. No. 0.0022;
Matches   143;   Conservative   0;   Mismatches   135;   Indels   2;   Gaps   1;

Qy 412 TGATGTTGACATTATTATTACTGCAAGATTAAATGTGGTTTACATATCTTTATGTACTGCCA 471
      ||||| |||||||||||||| ||| | ||| ||||||||| ||||| ||||||||||||
Db 1713 TTATGTAAGTATTTTTTTTTTATATGTAAGTTAGTTATTATTAAGTATTTGTTTTTTAGA 1772

Qy 472 TTTTTTGTTTCTGGTAAACTGGAATATAAAGTGAAAGAACAAACATTGAACATACTTAA 531
      || |||||| || || || ||| ||||| || |||||||| |||||| | |||||||
Db 1773 TGTTTTATATATAAAATATTCGTATATAATAATATGTTTGGAAATGTATATGTGTAATATATA 1832

Qy 532 TGTATTTTTA--TAGAACTTTGTAAACGAAGGAGATTCATGTTTTAGAAGTCTGTCCTT 589
      ||||||||| || ||||||| |||| |||||||||||| ||||| |||||||||||
Db 1833 TTTATATATGTATATGTGTAGCAGATATAAGATTAATTCGAACGTTGTATAAAATT 1892

Qy 590 TTTTATATCTTGAAAGAAAATCTATGTATGCTATAAAATAAATCCTATTATTTTTCT 649

```
     1893 TTTTTTATATTGAATAAAGTGTTTTTAGAATATATATATATATATATATATATATATAT 1952
          |||| |||| |||||| |||  ||  || |||| ||   |||| ||
      650 CAGGAATCGGTTAGGAATTGCAGGCAATGAGAGATTTTTG 689
          ||||||| ||||| |||||
     1953 ATATATACGGGTTATTTTTTTAATTGATGTGATGTTTAG 1992
```

SULT 6
-09-032-684-20
Sequence 20, Application US/09032684
Patent No. 5882874
GENERAL INFORMATION:
APPLICANT: Fisher, Paul B.
TITLE OF INVENTION: RECIPROCAL SUBTRACTION DIFFERENTIAL
TITLE OF INVENTION: DISPLAY
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESS:
ADDRESSEE: Cooper & Dunham LLP
STREET: 1185 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/032,684
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: White, John P.
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 55551/JPW/AMG
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 278-0400
TELEFAX: (212) 391-0525
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 177 base pairs
TYPE: nucleic acid
STRANDEDNESS: not relevant
TOPOLOGY: linear
MOLECULE TYPE: cDNA
S-09-032-684-20

```
Query Match       5.4%;    Score 45.2;  DB 2;  Length 177;
Best Local Similarity   60.5%;  Pred. No. 0.0096;
Matches   92;  Conservative   0;  Mismatches 58;  Indels  2;  Gaps  1;

Y   670 GCAGGCAATGAGAGATTTTTGCGGGGCAGGGATGGGAATGTTGTTCATAAATAATTAGAC 729
            ||||||||  |||||||||| |||| ||| |||||||||  |||
b     1 GTAGGCAATAAAAATGTTTTCAGAGGTGCGAAAAAGCTTTTGTTTTCTTAAACCATTCTTA 60

Y   730 ATTTTCTATAGATATTTGACATTCTGCGAAGCAACAGCAAACTGAAGACCAACTCCTA 789
            |||||||||||||||||||||||  |||||| ||||||||||||||||
b    61 GTCT--CTGCCACACTTGACACTCCGTCAAGTGAGAAGCGAACTAAAGACCAACTGCGG 118

Y   790 TGAGAAATATTATGATGTTTATGTAATAAAGA 821
            |||||||||||||||| ||| ||||
b   119 TGGAAAATATTATGTTTATGTAATAAAAAAAA 150
```

.ESULT 7
S-09-644-460-20
Sequence 20, Application US/09644460
Patent No. 6657053
GENERAL INFORMATION:
APPLICANT: Fisher, Paul B.
TITLE OF INVENTION: Reciprocal Subtraction Differential

; TITLE OF INVENTION: Display
; FILE REFERENCE: 34587-C-PCT-USA
; CURRENT APPLICATION NUMBER: US/09/644,460
; CURRENT FILING DATE: 2000-08-23
; PRIOR APPLICATION NUMBER: PCT/US99/04323
; PRIOR FILING DATE: 1999-02-26
; PRIOR APPLICATION NUMBER: US 09/197,889
; PRIOR FILING DATE: 1998-11-23
; PRIOR APPLICATION NUMBER: US 09/185,115
; PRIOR FILING DATE: 1998-11-03
; PRIOR APPLICATION NUMBER: US 09/032,684
; PRIOR FILING DATE: 1998-02-27
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 20
; LENGTH: 177
; TYPE: DNA
; ORGANISM: homo sapiens
US-09-644-460-20

```
Query Match       5.4%;    Score 45.2;  DB 4;  Length 177;
Best Local Similarity   60.5%;  Pred. No. 0.0096;
Matches   92;  Conservative   0;  Mismatches 58;  Indels  2;  Gaps  1;

Qy  670 GCAGGCAATGAGAGATTTTTGCGGGGCAGGGATGGGAATGTTGTTCATAAATAATTAGAC 729
            ||||||||  |||||||||| |||| ||| |||||||||  |||
Db    1 GTAGGCAATAAAAATGTTTTCAGAGGTGCGAAAAAGCTTTTGTTTTCTTAAACCATTCTTA 60

Qy  730 ATTTTCTATAGATATTTGACATTCTGCGAAGCAACAGCAAACTGAAGACCAACTCCTA 789
            |||||||||||||||||||||||  |||||| ||||||||||||||||
Db   61 GTCT--CTGCCACACTTGACACTCCGTCAAGTGAGAAGCGAACTAAAGACCAACTGCGG 118

Qy  790 TGAGAAATATTATGATGTTTATGTAATAAAGA 821
            |||||||||||||||| ||| ||||
Db  119 TGGAAAATATTATGTTTATGTAATAAAAAAAA 150
```

RESULT 8
US-09-621-976-2813
; Sequence 2813, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 2813
; LENGTH: 832
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 235..399
US-09-621-976-2813

```
Query Match       5.4%;    Score 45.2;  DB 4;  Length 832;
Best Local Similarity   24.8%;  Pred. No. 0.018;
Matches   74;  Conservative  88;  Mismatches 136;  Indels  0;  Gaps  0;

Qy  322 CTTTGAAGCATTGGTAGGAACTCTTAAAGCTGCAAAACGAAGGAAGATTGTAACACATCC 381
            |:|:::|||| ::::: :|:|: :|:|:::   ::   :|::::
Db  172 CWTKRWRASWYCWWWGKARKWSTWRKSRSYASARSAKRCCYSCSWGAMSWKYMWRMWRW 231

Qy  382 AGGAGAGCTGCTTCTGCAAGGTGTTCATGATGGTTGACATTATATTACTGCAAGATTA 441
            ::::::::::::::::::
Db  232 RGWATGAGMKAWRASCMMRRKYAGKSKTSYKSMWCWTRSWKYCYTKARWTGYYCYRKGG 291

Qy  442 ATGTGGTTTACATATCTTTATGTACTGCCATTTTTGTTTCTCGTAAACTGGAATATAAA 501
```

```
                    : :|:.:..|::|..:     :. : :|.|.:.:. : :..|:..|::::.:|.: |:..|:
292 MWGKRGRWYASKKYMWKRWWCWARMYRYSTGTRASMWWRRWRYYTMMMKWKYAWARRAW 351

502 GTCAAAGAACAAACATTTGAACATACTTAATGTATTTTATAGAACTTTGTAAACGAAAG 561
    : ::|.:..|::|                           |:::.|.|||||  |||:
352 RWAWMWAWRRACAAAATATATATTATTATGGTACAATTCTTGTACTTTAGCAAATCTGG 411

562 GAGATTCATGTTTTAGAAGTCTGTCCTTTTTTATATCTTGAAAGAAAATCTATGTATG 619
                    ||||||            ||||   |||||
412 AGTTAGTTCATATGTCAAAGTCAGTCGTTAATATTTCTTAGAGGAAAGTTTTGCTTTTTGTG 469
```

ULT 9
09-790-988-1/C
equence 1, Application US/09790988
atent No. 663295
ENERAL INFORMATION:
APPLICANT: SHIGENOBU, SHUJI
APPLICANT: WATANABE, HIDEMI
APPLICANT: HATTORI, MASAHIRA
APPLICANT: SAKAKI, YOSHIYUKI
TITLE OF INVENTION: GENOME DNA OF BACTERIAL SYMBIONT OF APHIDS
FILE REFERENCE: 081356/0159
CURRENT APPLICATION NUMBER: US/09/790,988
CURRENT FILING DATE: 2001-02-23
PRIOR APPLICATION NUMBER: JP2000-107160
PRIOR FILING DATE: 2000-04-07
NUMBER OF SEQ ID NOS: 7
SOFTWARE: PatentIn Ver. 2.1
EQ ID NO 1
  LENGTH: 640681
  TYPE: DNA
  ORGANISM: Buchnera sp.
-09-790-988-1

uery Match   5.4%;   Score 44.8;   DB 4;   Length 640681;
est Local Similarity 52.1%;   Pred. No. 0.34;
latches 125; Conservative   0; Mismatches 112; Indels 3; Gaps 1;

```
412 TGATGTTGACATTATTATTACTGCAAGATTAATGTGGTTACATATCTTTATGTACTGCCA 471
    ||||||||||||||||||||                          ||||||||
618831 TAATTTACATATATTCTTTTATTGTAACATTTGAATTTAAAATATATTTTCATGTTTTCTAT 618772

472 TTTTTGTTTCTGGTAAACTGGAATATAAAGTGAAAGAACAAACATTTGAACATACTTAA 531
    ||||||||                    ||||||         ||||||
618771 TTTTTTTATTTTTGAAAATTAATAATTGACTTTTAAAAAAAAACTAATTAATAATTTACAT 618712

532 TGTATTTTATAGAACTTTGTAAACGAAAGGAGATTCATGTTTAGAAGTCTGTCCTTTT 591
    ||||||                   |||||||                ||||
618711 TGACAATTTTTTATATATTTTAAGTTCAAAAAA--ATATTTTATAAAAATATACTAAA 618655

592 TTATATCTTGAAAGAAAATCATGTTATGATGCTATAAAATAAATCCTATTATTTTTCTCA 651
                     |||                   |||
618654 AAAATAATTTACAATAGGATCAAAATATTATGTTACACCAATTTATTATTAAATCCCAATCA 618595
```

SULT 10
-08-731-722-4
Sequence 4, Application US/08731722
Patent No. 5961971
GENERAL INFORMATION:
  APPLICANT: Martin, Frank N.
  TITLE OF INVENTION: Biocontrol of Fungal Soilborne Pathogens
  TITLE OF INVENTION: by Pythium oligandrum
  NUMBER OF SEQUENCES: 8
  CORRESPONDENCE ADDRESS:
    ADDRESSEE: Saliwanchik & Saliwanchik
    STREET: 2421 N.W. 41st Street, Suite A-1
    CITY: Gainesville
    STATE: FL
    COUNTRY: US
    ZIP: 32606-6669
  COMPUTER READABLE FORM:
    MEDIUM TYPE: Floppy disk
    COMPUTER: IBM PC compatible
    OPERATING SYSTEM: PC-DOS/MS-DOS
    SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
  APPLICATION NUMBER: US/08/731,722
  FILING DATE:
  CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
  NAME: Whitlock, Ted W.
  REGISTRATION NUMBER: 36,965
  REFERENCE/DOCKET NUMBER: UF-161
TELECOMMUNICATION INFORMATION:
  TELEPHONE: 352-375-8100
  TELEFAX: 352-372-5800
INFORMATION FOR SEQ ID NO: 4:
  SEQUENCE CHARACTERISTICS:
    LENGTH: 1218 base pairs
    TYPE: nucleic acid
    STRANDEDNESS: single
    TOPOLOGY: linear
  MOLECULE TYPE: DNA (genomic)
  ORIGINAL SOURCE:
    INDIVIDUAL ISOLATE: 17-1
US-08-731-722-4

Query Match   5.3%;   Score 44.4;   DB 2;   Length 1218;
Best Local Similarity 49.1%;   Pred. No. 0.034;
Matches 172; Conservative   0; Mismatches 176; Indels 2; Gaps 2;

```
Qy 486 TAAACTGGAATATAAAGTGAAAGAACAAACATTTGAACATACTTAATGTATTTTTATAGA 545
        ||||||||||                    |||||||||      |   ||||||
Db 440 TRACTTTTAAAAAAAATTAAAATGAAAATACTTCAATAAGAATATATTTTATTATAAA 499

Qy 546 ACTTTGTAAACGAAGGAGATTCATGTTTTAGAAGTCGTCCTTTTTATATCTTGAAAG 605
         |||      ||||||                 ||||||
Db 500 TATTAATTAATTTGTGTAAATCCACCAGAAAATTTAACTATTATTAAATATTG-AAG 558

Qy 606 AAAATCTATGTATGATGCTATAAAATAAAATCCATTATTTTCTCAGGAATCTGGTTAGG 665
         ||||                          |||||      |||||||||||
Db 559 AACTTTTACT-ATAT-TTCTAAATAATAATCTGAATTTAAACAGTAGACCCTATGAACCTT 618

Qy 666 AATTGCAGGCAATGAACATGAGATTTTTTGCGGGGCAGGGATGGGAAATGTTTGTTGTTCATAAATAATT 725
         |||||                    |||||||||||
Db 619 TAATAATAGCTATAAGAAATTTTCAATAGAGAAAGGTTTTATTTTTTAATGAAAATT 678

Qy 726 AGACATTTTCTATAGATATTTGACATTCTGCGAAAGCAACAAGCAAACTGAAGACCAACT 785
                        ||||||||||                      |||||
Db 679 GTAATTCTTATTTACTTTAATAAGAATAATAATTATTATAATAAA-ATATATT 737

Qy 786 CCTATCAGAAATATTATGATGTTTATGTAATAAAGACATGTAACTGTCTT 835
         |||||                 |||||||||||
Db 738 CTTATTAGAAGTATTTTCATTTTAATTTTTTTTAAAAGTTATATATCTT 787
```

RESULT 11
US-10-204-708-42
; Sequence 42, Application US/10204708
; Patent No. 667731
; GENERAL INFORMATION:
;   APPLICANT: OLEK, Alexander
;   APPLICANT: PIEPENBROCK, Christian
;   APPLICANT: BERLIN, Kurt
;   TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA Replication
;   TITLE OF INVENTION: by Assessing DNA Methylation
;   FILE REFERENCE: 5013.1012
;   CURRENT APPLICATION NUMBER: US/10/204,708
;   CURRENT FILING DATE: 2003-05-06
;   PRIOR APPLICATION NUMBER: PCT/EP01/03971
;   PRIOR FILING DATE: 2001-04-06
;   PRIOR APPLICATION NUMBER: DE 10019058.8
;   PRIOR FILING DATE: 2000-04-06
;   PRIOR APPLICATION NUMBER: DE 10019173.8

PRIOR FILING DATE: 2000-04-07
PRIOR APPLICATION NUMBER: DE 10032529.7
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: DE 10043826.1
PRIOR FILING DATE: 2000-09-01
NUMBER OF SEQ ID NOS: 98
SEQ ID NO 42
LENGTH: 8537
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
3-10-204-708-42

Query Match   5.2%;   Score 43.8; DB 4; Length 8537;
Best Local Similarity 51.0%; Pred. No. 0.11; Indels 1; Gaps 1;
Matches 128; Conservative 0; Mismatches 122; Indels 1; Gaps 1;

```
517  TTGAACACATACTTAATGTATTTTTATAGAACTTGTAAACGAAAGGAGATTCATGTTTTA 576
     ||||  ||  ||| |||||||||||||||||||||  || ||||||||| |||||||||
7358 TTTTTTATTTTTTTTTTTTTTTTTTAAAAAAAAAATTTGAATTTTAAGGATTAATGTTTTT 7417

577  GAAGTCTGTCCTTTTTTATATCTTGAAAGAAAATCATGTATGATGCTATAAAATAAATC 636
     |    *          | |    |||  |||
7418 GAGTTATGAGAGATATTTT-TATATTTTAATAATTTATTTTTATTTTTATTGATTAAATT 7476

637  CTATTATTTTTCTCAGGAATCTCGGTTAGGGAATTGCAGGCAATGAGAGATTTTTTGCGGGGCA 696
     ||||  ||  ||    |||  ||||||||
7477 AGTTTGAGTAGTTTTGTTTTTTTTTTTAAATAATTATTGAGAATTTGGAAATAAGG 7536

697  GGGATGGGAATGTTGTTCATAAAATAATTAGACATTTCTATAGATATTTGACATTCTGC 756
     |||||         |||    ||||  |||  |
7537 TAAATTAGATTTTTTTTTTAATAGTTAAATTGGTATTTTTGAGAATATTTATAAATGTGT 7596

757  GAAAGCAACAA 767
     |||||||||
7597 TAATTGAAAAA 7607
```

ESULT 12
US-09-221-017B-77/C
Sequence 77, Application US/09221017B
Patent No. 6444799
GENERAL INFORMATION:
APPLICANT: Ross, Bruce C.
TITLE OF INVENTION: P. GINGIVALIS NUCLEOTIDES AND USES THEREOF
NUMBER OF SEQUENCES: 1120
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORRISON & FOERSTER
STREET: 755 PAGE MILL ROAD
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304-1018
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows
SOFTWARE: FastSEQ for Windows Version 2.0b
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/221,017B
FILING DATE: 23-DEC-1998
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PP1182
FILING DATE: 31-DEC-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PP1546
FILING DATE: 30-JAN-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PP2911
FILING DATE: 09-APR-1998
PRIOR APPLICATION DATA:

APPLICATION NUMBER: PCT/AU98/01023
FILING DATE: 10-DEC-1998
ATTORNEY/AGENT INFORMATION:
NAME: Monroy, Gladys H
REGISTRATION NUMBER: 32,430
REFERENCE/DOCKET NUMBER: 27340-20021.00
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-813-5600
TELEFAX: 650-494-0792
TELEX: 706141
INFORMATION FOR SEQ ID NO: 77:
SEQUENCE CHARACTERISTICS:
LENGTH: 1189 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: circular
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: UNKNOWN
ORIGINAL SOURCE:
ORGANISM: PORYPHYROMONAS GINGIVALIS
FEATURE:
NAME/KEY: misc_feature
LOCATION: 1...1189
US-09-221-017B-77

Query Match   5.1%;   Score 43; DB 4; Length 1189;
Best Local Similarity 51.3%; Pred. No. 0.078; Indels 0; Gaps 0;
Matches 100; Conservative 0; Mismatches 95; Indels 0; Gaps 0;

```
Qy 420 ACATTATATTACTGCAAGATAATGTGGTTTACATATCTTTATGTACTGCCATTTTTTGT 479
        |  |||| ||  ||  |   | | ||||
Db 485 ACTTTATTATTGATGCCATGTCAAGTGTGTTTCATACATAACTGTACAACAGGGTGATGA 426

Qy 480 TTCTGGTAAACTGGAATATAAGTGAAAGAACAAACATTGAACATACTTAATGTATTTT 539
        ||  |||    ||  || |  ||||||  ||
Db 425 TTATAAGAAACAGAATGATAATGAGAGATTTTACCGGCCAGTTGTATATGATGAGAATGTTGT 366

Qy 540 TATAGAACTTGTAAACGAAAGGAGATTCATGTTTTTAGAAGTCTGTCCTTTTTTATATCT 599
        ||    |||  |  |||    ||  |||
Db 365 TGCAATACTTAGAAAATAAAAGGGTTCATCAGAGAGGATGATGGACATCATCTGTTACA 306

Qy 600 TGAAAGAAAATCTAT 614
        ||||| || | |
Db 305 TCGAAATAATGATAT 291
```

RESULT 13
US-08-239-953-1/C
; Sequence 1, Application US/08299953
; Patent No. 5646333
GENERAL INFORMATION:
APPLICANT: Dobres, Michael S. and Mandaci, Sevnur
TITLE OF INVENTION: A Plant Promoter Useful for Directing the
TITLE OF INVENTION: Expression of Foreign Proteins to the Plant Epidermis
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. 5646333ris
STREET: One Liberty Place 46th. Floor
CITY: Philadelphia
STATE: PA
ZIP: 19103
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/299,953
FILING DATE: Herewith
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Beardell, Lori Y.

Best Local Similarity 48.7%; Pred. No. 0.13;
Matches 113; Conservative   1; Mismatches 118; Indels   0; Gaps   0;

```
403 TGTTCATGATGATGTTGACATTATATTACTGCAAGATTAATGTGGTTACATATCTTTAT 462
        ||||  ||| || |||||  ||| ||||||| | || |  ||||| |  | ||||
826 TGTTAGTAATAATGGTGATATGCCAGTATTAAATCAATATTAAAATATTAATTTATATCA 767

463 GTACTGCCATTTTTGTTTCTGGTAAACTGGAATATAAAGTGAAAAGAACAAACATTTGAA 522
         |||  ||| |||| ::|   |  ||| |||| ||| ||  ||||  || |||||||
766 TTTTTTAGATTAATTTTTTATAAAAASTTAATCATTTATTTTAAACAAATAAATGTTTAA 707

523 CATACTTAATGTATTTTATAGAACTTTGTAAACGAAAGGAGATTCATGTTTTTAGAAGTC 582
          |  ||||| |   |  |||   |   | ||| || | ||  | |||  ||||| |
706 TAAATATCACATTAAATGTTTTTTCTTCTTAGTCTGATCTTTAAAAAATATTTGGCATAA 647

583 TGTCCTTTTTTATATCTTGAAAGAAAATCATGTATGATGCTATAAAATAAA 634
        ||||||| ||  |   |  ||| ||  ||  ||  |   ||||  |||
646 TTTTTTTTTTAATATCTATAACAAATTTTTATAAAGAGTGATGTTATATTAA 595
```

arch completed: March 17, 2004, 14:25:20
)b time : 95 secs

```
                    Copyright  GenCore version 5.1.6
                    Copyright (c) 1993 - 2004 Compugen Ltd.

nucleic - nucleic search, using sw model

on:     March 17, 2004, 12:27:16 ; Search time 379 Seconds
                                    (without alignments)
                                    8111.172 Million cell updates/sec

le:         US-09-648-310-3
fect score: 835
quence:     1 ggcacgaggcttgagcgcag........taagacatgtaactgtctt 835

 oring table: IDENTITY_NUC
             Gapop 10.0 , Gapext 1.0

arched:     242557 seqs, 1840798884 residues

tal number of hits satisfying chosen parameters:     4865114

nimum DB seq length: 0
ximum DB seq length: 2000000000

st-processing: Minimum Match 0%
               Maximum Match 100%
               Listing first 45 summaries

                                    SUMMARIES

cabase :    Published_Applications_NA.*
```

sult No.	Score	% Query Match	Length	DB	ID	Description
1	619.4	74.2	778	9	US-09-925-300-545	Sequence 545, App
2	542.8	65.0	552	9	US-09-920-300A-939	Sequence 939, App
3	542.8	65.0	552	13	US-10-033-528-939	Sequence 939, App
4	542.8	65.0	552	14	US-10-099-926-939	Sequence 939, App
5	469.8	56.3	483	12	US-10-085-783A-56189	Sequence 56189, A
6	469.8	56.3	483	15	US-10-242-535A-56189	Sequence 56189, A
7	382.4	45.8	406	9	US-09-880-107-687	Sequence 687, App
8	381	45.6	400	12	US-10-085-783A-21733	Sequence 21733, A
9	381	45.6	400	15	US-10-242-535A-21733	Sequence 21733, A
10	364	43.6	365	9	US-09-878-178-2101	Sequence 2101, Ap
11	364	43.6	365	13	US-10-046-935-2101	Sequence 2101, Ap
12	364	43.6	365	14	US-10-146-502-2101	Sequence 2101, Ap
13	237	28.4	458	10	US-09-911-995-26075	Sequence 26075, A
14	188.4	22.6	388	9	US-09-925-299-318	Sequence 318, App
15	188.4	22.6	388	10	US-09-925-299-318	Sequence 318, App

```
Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.
```

```
16  170.2  20.4  611         10  US-09-871-161-376            Sequence 376, App
17  109.4  13.1  486          9  US-09-938-842A-355           Sequence 355, App
18  109.4  13.1  486         11  US-09-938-842A-355           Sequence 355, App
c   19  109.4  13.1  492     10  US-09-770-961-618            Sequence 618, App
20  98.6   11.8  533         12  US-10-425-114-19867          Sequence 19867, A
c   21  91     10.9  716     12  US-10-424-599-94878          Sequence 94878, A
22  69.8   8.4   268          9  US-09-294-093B-2020          Sequence 2020, Ap
23  60     7.2   60          10  US-09-908-975-22102          Sequence 22102, A
24  55     6.6   3673778     14  US-10-312-841-1              Sequence 1, Appli
25  54.2   6.5   3673778     14  US-10-312-841-2              Sequence 2, Appli
26  52.2   6.3   113515      14  US-10-311-455-2147           Sequence 2147, Ap
27  50.6   6.1   6073        14  US-10-311-455-1515           Sequence 1515, Ap
28  50     6.0   50          15  US-10-131-827-722            Sequence 722, App
29  50     6.0   6020        14  US-10-204-708-8              Sequence 8, Appli
30  50     6.0   6020        14  US-10-311-455-172            Sequence 172, App
c   31  49.4  5.9  5467      14  US-10-311-455-2082           Sequence 2082, Ap
32  49.4   5.9   14950       14  US-10-311-455-1230           Sequence 1230, Ap
33  49.2   5.9   5488        14  US-10-311-455-1429           Sequence 1429, Ap
34  49.2   5.9   6381        14  US-10-311-455-939            Sequence 939, App
35  49.2   5.9   6381        14  US-10-240-485-71             Sequence 71, Appl
36  49     5.9   6210        14  US-10-240-453-346            Sequence 346, Appl
c   37  48.8  5.8  5728      14  US-10-311-455-74             Sequence 74, Appl
38  48.4   5.8   13038       14  US-10-311-455-1248           Sequence 1248, Ap
39  48.2   5.8   5937        14  US-10-240-485-96             Sequence 96, Appl
40  48     5.7   6161        14  US-10-311-455-596            Sequence 596, App
41  47.6   5.7   6361        14  US-10-311-455-1113           Sequence 1113, Ap
42  47.4   5.7   6412        12  US-10-221-613-104            Sequence 104, App
43  46.8   5.6   972         12  US-10-424-599-93099          Sequence 93099, A
44  46.8   5.6   13784       16  US-10-257-166-144            Sequence 144, App
45  46.6   5.6   5324        14  US-10-311-455-1763           Sequence 1763, Ap
```

                                    ALIGNMENTS

```
RESULT 1
US-09-925-300-545
; Sequence 545, Application US/09925300
; Patent No. US2002015168A1
; GENERAL INFORMATION:
;  APPLICANT: Craig Rosen,
;  APPLICANT: Steve Ruben
;  TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
;  FILE REFERENCE: PA101
;  CURRENT APPLICATION NUMBER: US/09/925,300
;  CURRENT FILING DATE: 2001-08-10
;  PRIOR APPLICATION NUMBER: PCT/US00/05988
;  PRIOR FILING DATE: 2000-03-08
;  PRIOR APPLICATION NUMBER: 60/124,270
;  PRIOR FILING DATE: 1999-03-12
;  NUMBER OF SEQ ID NOS: 1890
;  SOFTWARE: PatentIn Ver. 2.0
;  SEQ ID NO 545
;  LENGTH: 778
;  TYPE: DNA
;  ORGANISM: Homo sapiens
;  FEATURE:
;  NAME/KEY: misc feature
;  LOCATION: (641)
;  OTHER INFORMATION: n equals a,t,g, or c
;  NAME/KEY: misc feature
;  LOCATION: (652)
;  OTHER INFORMATION: n equals a,t,g, or c
US-09-925-300-545

Query Match             74.2%;  Score 619.4;  DB 9;  Length 778;
Best Local Similarity   99.8%;  Pred. No. 1.1e-141;
Matches  620;  Conservative    0;  Mismatches   1;  Indels    0;  Gaps    0;

Qy     26  CTTACTTTTCCCCTACCCTGCTCCTCCTCCACAGCCGTCTTTCTCTTTGCCTCAGC  85
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     15  CGTACTTTTCCCCTACCCTGCTCCTCCTCCACAGCCGTCTTTCTCTTTGCCTCAGC  74
```

RESULT 2 (left column)

```
Qy   86  CACTTCCTTCCTTCGCCTCACCCTCCCAGTGCACTGAAGAAGGTAACCGGTCCAGACC  145
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   75  CACTTCCTTCCTTCGCCTCACCCTCCCAGTGCACTGAAGAAGGTAACCGGTCCAGACC  134

Qy  146  CACGCGGCGCCAGTTCTCCGGCGGGGAAGGGAAAACCGCGCAGAGAGGCAGCAATGAATGTG  205
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  135  CACGCGGCGCCAGTTCTCCGGCGGGGAAGGGAAAACCGCGCAGAGAGGCAGCAATGAATGTG  194

Qy  206  GATCACGAGGTTAACCTCTTAGTGGAGGAGAAATTCATCGTTTGGGTTCAAAAAAATGCTGAT  265
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  195  GATCACGAGGTTAACCTCTTAGTGGAGGAGAAATTCATCGTTTGGGTTCAAAAAAATGCTGAT  254

Qy  266  GGAAAGTTAAGCGTGAAATTGGGGTCCTCTCCGTGATGAAATGTGCCAACCTCTT  325
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  255  GGAAAGTTAAGCGTGAAATTGGGGTCCTCTCCGTGATGAAATGTGCCAACCTCTT  314

Qy  326  GAAGCATTGGTAGGAACTCTTAAAGCTGCAAAACGAAGGAAGGAAGTTGAACATATCCAGGA  385
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  315  GAAGCATTGGTAGGAACTCTTAAAGCTGCAAAACGAAGGAAGGAAGTTGAACATATCCAGGA  374

Qy  386  GAGCTGCTTCTGCAAGGTGTTCATGATGTTGACATTATTACTGCAAGATTAATGT  445
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  375  GAGCTGCTTCTGCAAGGTGTTCATGATGTTGACATTATTACTGCAAGATTAATGT  434

Qy  446  GGTTTACATATCTTTATGTACTGCCATTTTTGTTCTGGTAAACTGGAATATAAAGTGA  505
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  435  GGTTTACATATCTTTATGTACTGCCATTTTTGTTCTGGTAAACTGGAATATAAAGTGA  494

Qy  506  AAGAACAAACATTTGAACATTGAACATTAATGTATTTTATACTTAAGCAGAAGGAGA  565
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  495  AAGAACAAACATTTGAACATTGAACATTAATGTATTTTATACTTAAGCAGAAGGAGA  554

Qy  566  TTCATGTTTTAGAAGTCTGTCCTTTTTATATCCTATTATTT  646
         |||||||||||||||||||||||||||||||||||||||||
Db  615  TTCATGTTTTAGAAGTCTGTCCTTTTTATATCCTATTATTT  635
```

```
Qy  290  GTCCTCTTCCGTGATGATGAATAAATGTGCCAACCTCTTTGAAGCATTGGTAGGAACTCTTAAA  349
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  552  GTCCTCTTCCGTGATGATGAATAAATGTGCCAACCTCTTTGAAGCATTGGTAGGAACTCTTAAA  493

Qy  350  GCTGCAAAACGAAGGAAGGAAGATTGTAACATATCCAGGAGAGCTGCTTCTGCAAGGTGTTCAT  409
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

RESULT 3 (right column)

```
Db  492  GCTGCAAAACGAAGGAAGGAAGATTGTAACATATCCAGGAGAGCTGCTTCTGCAAGGTGTTCAT  433

Qy  410  GATGATGTTGACATTATATTACTGCAAGATTAATGTGGTTTACATATCTTTATGTACTGC  469
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  432  GATGATGTTGACATTATATTACTGCAAGATTAATGTGGTTTACATATCTTTATGTACTGC  373

Qy  470  CATTTTTGTTCTGGTAAACTGGAATATAAAGTGAAAGACAAACAACATTGAACATACTT  529
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  372  CATTTTTGTTCTGGTAAACTGGAATATAAAGTGAAAGACAAACAACATTGAACATACTT  313

Qy  530  AATGTATTTTATAGAACTTGTAAACGAAAGGAGATTCATGTTTTAGAAGTCTGTCCTT  589
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  312  AATGTATTTTATAGAACTTGTAAACGAAAGGAGATTCATGTTTTAGAAGCCTGTCCTT  253

Qy  590  TTTATATCTTGAAAGAAAATCATGTATGATGCGTATAAAAATAAAATCCTATTATTTTCT  649
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  252  TTTATATCTTGAAAGAAAATCATGTATGATGCGTATAAAAATAAAATCCTATTATTTTCT  193

Qy  650  CAGGAATCTGGTTAGGAATTGCAGGCAATGACGAGATTTTTGCGGGGCAGGGATGGGAATGT  709
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  192  CAGGAATCTGGTTAGGAATTGCAGGCAATGACGAGATTTTTGCGGGGCAAGGATGGGAATGT  133

Qy  710  TTGTTCATAAATAATTAGACATTTCTATAGATATTTGACATTGAACATTCTGCGAAAGCAACAAGC  769
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  132  TTGTTCATAAATAATTAGACATTTCTATAGATATTTGACATTGAACATTCTGCGAAAGCAACAAGC  73

Qy  770  AAACTGAAGACCAACTCCTATGAGAAAATATTATGATGTTTATGTAATAAAGACATGTAAC  829
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   72  AAACTGAAGACCAACTCCTATGAGAAAATATTATGATGTTTATGTAATAAAGACATGTAAC  13

Qy  830  TGTCTT  835
         ||||||
Db   12  TGTCTT  7
```

```
Qy  290  GTCCTCTTCCGTGATGATGAATAAATGTGCCAACCTCTTTGAAGCATTGGTAGGAACTCTTAAA  349
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  552  GTCCTCTTCCGTGATGATGAATAAATGTGCCAACCTCTTTGAAGCATTGGTAGGAACTCTTAAA  493

Qy  350  GCTGCAAAACGAAGGAAGGAAGATTGTAACATATCCAGGAGAGCTGCTTCTGCAAGGTGTTCAT  409
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  492  GCTGCAAAACGAAGGAAGGAAGATTGTAACATATCCAGGAGAGCTGCTTCTGCAAGGTGTTCAT  433

Qy  410  GATGATGTTGACATTATATTACTGCAAGATTAATGTGGTTTACATATCTTTATGTACTGC  469
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  432  GATGATGTTGACATTATATTACTGCAAGATTAATGTGGTTTACATATCTTTATGTACTGC  373
```

Db   312 AATGTATTTTTATAGACACTTGTAAACGAAAGGAGATTCATGTTTTGAGAGCCTGTCCTT 253

Qy   590 TTTTATATCTTGAAAGAAAATCTATGATGATGTGATGATGCTATAAAATAAATCCTATTATTTTTCT 649
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   252 TTTTATATCTTGAAAGAAAATCTATGTATGATGCTATAAAATAAATCCTATTATTTTTCT 193

Qy   650 CAGGAATCTGGTTAGGAATTGCAGGCAATGCAATGAGATTTTTGCGGGGCAGGGATGGGAATGT 709
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   192 CAGGAATCTGGTTAGGAATTGCAGGCAATGCAATGAGATTTTTGCGGGGCAAGGATGGGAATGT 133

Qy   710 TTGTTCATAAATAATTAGACATTTTCTATAGATATTGACATTCTGCGAAAGCAACAAGC 769
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   132 TTGTTCATAAATAATTAGACATTTTCTATAGATATTGACATTCTGCGAAAGCAACAAGC 73

Qy   770 AAACTGAAGACCAACTCCTATGAGAAATATTATGTGTAATAAAGACATGTAAC 829
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   72 AAACTGAAGACCAACTCCTATGAGAAATAT ATGATGTTATGTAATAAAGACATGTAAC 13

Qy   830 TGTCTT 835
          ||||||
Db   12 TGTCTT 7

RESULT 5
US-10-085-783A-56189
; Sequence 56189, Application US/10085783A
; Publication No. US20040037841A1
; GENERAL INFORMATION:
;   APPLICANT: Chondrogene Inc.
;   APPLICANT: Liew, C.C.
;   TITLE OF INVENTION: Compositions and Methods Relatting to Osteoarthritis
;   FILE REFERENCE: 4231/2002
;   CURRENT APPLICATION NUMBER: US/10/085,783A
;   CURRENT FILING DATE: 2002-02-28
;   PRIOR APPLICATION NUMBER: US 60/305,340
;   PRIOR FILING DATE: 2001-07-13
;   PRIOR APPLICATION NUMBER: US 60/275,017
;   PRIOR FILING DATE: 2001-03-12
;   PRIOR APPLICATION NUMBER: US 60/271,955
;   PRIOR FILING DATE: 2001-02-28
;   NUMBER OF SEQ ID NOS: 58994
;   SOFTWARE: PatentIn version 3.2
;   SEQ ID NO 56189
;   LENGTH: 483
;   TYPE: DNA
;   ORGANISM: Human
US-10-085-783A-56189

Query Match   56.3%;   Score 469.8;   DB 12;   Length 483;
Best Local Similarity   99.6%;   Pred. No. 4.3e-105;
Matches 471;   Conservative   0;   Mismatches   2;   Indels   0;   Gaps   0;

Qy   127 AGGTAACCGGGTCCAGACCCACGCGCGCCAGTTCCCGGCGGGAAGGAAAACCCGCGCAG 186
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   11 AAGTAACTGGGTCCAGACCCACGCGCGCCAGTTCCCGGCGGGAAGGAAAATCATCGTT 70

Qy   187 AGAGGCAGCAATGAATGGATCACGAGGTTAACCTCTTAGTGGAGGAAATTCATCGTT 246
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   71 AGAGGCAGCAATGAATGGATCACGAGGTTAACCTCTTAGTGGAGGAAATTCATCGTT 130

Qy   247 GGGTTCAAAAATGCTGATGAAAGTTAAGCTGAAATTTGGGGTCCTCTTCCGTGATGA 306
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   131 GGGTTCAAAAAATGCTGATGAAAGTTAAGCGTGAAATTTCGGGTCCTCTTCCGTGATGA 190

Qy   307 TAAATGTGCCAACCTCTTGAAGCATGGTAGGAACTCTTAAAGCTGCAAAACGAAGGAA 366
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   191 TAAATGTGCCAACCTCTTGAAGCATTGGTAGGAACTCTTAAAGCTGCAAAACGAAGGAA 250

Qy   367 GATTGTAACATATCCAGGAGAGCTGCTTCATGATGATGTTCATGATGTTGACATTAT 426
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   251 GATTGTAACATATCCAGGAGAGCTGCTTCTGCAAGGTGTTCATGATGATGTTGACATTAT 310

Qy   470 CATTTTTTGTTCTCTGGTAAACTGGTAAACATAAAGTGAAAGAACAAACATTTGAACATACTT 529
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   372 CATTTTTTGTTCTCTGGTAAACTGGTAAACATAAAGTGAAAGAACAAACATTTGAACATACTT 313

Qy   530 AATGTATTTTTATAGACACTTGTAAACGAAAGAGATTCATGTTTCTCTGAGTCTGTCCTT 589
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   312 AATGTATTTTTATAGAACTTGTAAACGAAAGGAGATTCATGTTTTGAGAGCCTGTCCTT 253

Qy   590 TTTTATATTTTATAGAAAGAAAATCTATGTATGATGCTATAAAATAAATCCTATTATTTTTCT 649
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   252 TTTTATATCTTCTGTAAACGAAAGAGATTCATGTTTCTCTGAGTCTGTCCTT 193

Qy   650 CAGGAATCTGGTTAGGAATTGCAGGCAATGAGATTTTTTGCGGGGCAGGGATGGGAATGT 709
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   192 CAGGAATCTGGTTAGGAATTGCAGGCAATGAGATTTTTTGCGGGGCAAGGATGGGAATGT 133

Qy   710 TTGTTCATAAATAATTAGACATTTTCTATAGATATTATTGATGATGTTATGTAATAAAGACATGTAAC 769
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   132 TTGTTCATAAATAATTAGACATTTTCTATAGATATTGATGATGTTTATGTAATAAAGACATGTAAC 73

Qy   770 AAACTGAAGACCAACTCCTATGAGAAATATTATTATGATGTTATGTAATAAAGACATGTAAC 829
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   72 AAACTGAAGACCAACTCCTATGAGAAATATTATTGAACATTCTGCGAAAGCAACATGTAAC 13

Qy   830 TGTCTT 835
          ||||||
Db   12 TGTCTT 7

SULT 4
-10-099-926-939/C
Sequence 939, Application US/10099926
Publication No. US20030166064A1
GENERAL INFORMATION:
   APPLICANT: King, Gordon E
   APPLICANT: Meagher, Madeleine Joy
   APPLICANT: Xu, Jiangchun
   APPLICANT: Secrist, Heather
   APPLICANT: Jiang, Yuqiu
   TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
   TITLE OF INVENTION: AND DIAGNOSIS OF COLON CANCER
   FILE REFERENCE: 210121.547C2
   CURRENT APPLICATION NUMBER: US/10/099,926
   CURRENT FILING DATE: 2002-03-17
   NUMBER OF SEQ ID NOS: 1982
   SOFTWARE: FastSEQ for Windows Version 4.0
   SEQ ID NO 939
   LENGTH: 552
   TYPE: DNA
   ORGANISM: Homo sapiens
-10-099-926-939

Query Match   65.0%;   Score 542.8;   DB 14;   Length 552;
Best Local Similarity   99.6%;   Pred. No. 5.7e-123;
Matches 544;   Conservative   0;   Mismatches   2;   Indels   0;   Gaps   0;

,   290 GTCCTCTTCCGTGATGATGAAATGTGCCAACCTCTTTGAAGCATTGGTAGGAACTCTTAAA 349
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
     552 GTCCTCTTCCGTGATGATGAAATGTGCCAACCTCTTTGAAGCATTGGTAGGAACTCTTAAA 493

,   350 GCTGCAAAACGAAGGAAGAGATTGTAACATATCCAGGAGAGCTGCTTCTGCAAGGTGTTCAT 409
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
     492 GCTGCAAAACGAAGGAAGAGATTGTTAACATATCCAGGAGAGCTGCTTCTGCAAGGTGTTCAT 433

,   410 GATGATGTTGACATTAITATACTGCAAGATTAATGTGGTTACATATCTTTATGTACTGC 469
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
     432 GATGATGTTGACATTATATTACTGCAAGATTAATGTGGTTTACATATCTTTATGTACTGC 373

,   470 CATTTTTTGTTCTCTGGTAAACTGGTAAACTGGTAAATATAAAGTGAAAGAACAAACATTTGAACATACTT 529
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
     372 CATTTTTTGTTCTCTGGTAAACTGGTAAACTGGTAAATATAAAGTGAAAGAACAAACATTTGAACATACTT 313

>   530 AATGTATTTTTATAGACACTTGTAAACGAAAGAGATTCATGTTTCTCTGAGTCTGTCCTT 589
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

427 ATTACTGCAAGATTAATGTGGTTACATATCTTATGTGACTGCCATTTTTGTTCTGGT 486
    |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
311 ATTACTGCAAGATTAATGTGGTTACATATCTTATGTGACTGCCATTTTTGTTCTGGT 370

487 AAACTGGAATATAAAGTGAAAGAACAAACATTTGAACATACTTAATGTATTTTATAGAA 546
    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
371 AAACTGGAATATAAAGTGAAAGAACAAACATTTGAACATACTTAATGTATTTTATAGAA 430

547 CTTTGTAAACGAAAGGAGATTCATGTTTTAGAAGTCTGTCCTTTTTTATATCT 599
    |||||||||||||||||||||||||||||||||||||||||||||||||||||
431 CTTTGTAAACGAAAGGAGATTCATGTTTTAGAAGTCTGTCCTTTTTTATATCT 483

RESULT 6
S-10-242-535A-56189
; Sequence 56189, Application US/10242535A
; Publication No. US20040013663A1
; GENERAL INFORMATION:
;   APPLICANT: ChondroGene Inc.
;   APPLICANT: Liew, C.C.
;   TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis
;   FILE REFERENCE: 4231/2005
;   CURRENT APPLICATION NUMBER: US/10/242,535A
;   CURRENT FILING DATE: 2002-09-12
;   PRIOR APPLICATION NUMBER: US 10/085,783
;   PRIOR FILING DATE: 2002-02-28
;   PRIOR APPLICATION NUMBER: US 60/305,340
;   PRIOR FILING DATE: 2001-07-13
;   PRIOR APPLICATION NUMBER: US 60/275,017
;   PRIOR FILING DATE: 2001-03-12
;   PRIOR APPLICATION NUMBER: US 60/271,955
;   PRIOR FILING DATE: 2001-02-28
;   NUMBER OF SEQ ID NOS: 58994
;   SOFTWARE: PatentIn version 3.2
; SEQ ID NO 56189
;   LENGTH: 483
;   TYPE: DNA
;   ORGANISM: Human
S-10-242-535A-56189

Query Match   56.3%; Score 469.8; DB 15; Length 483;
Best Local Similarity   99.6%; Pred. No. 4.3e-105;
Matches   471; Conservative   0; Mismatches   2; Indels   0; Gaps   0;

Qy 127 AGGTAACCGGGTCCAGACCCACCGCGGCGCCCAGTTCTTCCGGCGGGAAGGAAAACCGCGCAG 186
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   11 AGGTAACTGGGTCCAGACCCACCGCGGCGCCCAGTTCTTCCGGCGGGAAGGAAAACCGCGCAG 70

Qy 187 AGAGGCAGCAATGAATGTGGATCACGAGGTTAACCTCTTAGTGGAGGAAATTCATCGTTT 246
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   71 AGAGGCAGCAATGAATGTGGATCACGAGGTTAACCTCTTAGTGGAGGAAATTCATCGTTT 130

Qy 247 GGGTCAAAAATGCTGATGGAAGTTAAGCGTGAAATTGGGGTCCTCTTCCGTGATGA 306
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 131 GGGTCAAAAATCGTGATGGAAGTTAAGCGTGAAATTGGGGTCCTCTTCCGTGATGA 190

Qy 307 TAAATGTGCCAACCTCTTTGAAGCATTGGTAGGAACTCTTAAAGCTGCAAAACGAAGGAA 366
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 191 TAAATGTGCCAACCTCTTTGAAGCATTGGTAGGAACTCTTAAAGCTGCAAAACGAAGGAA 250

Qy 367 GATTGTAACATATCCAGGAGAGCTGCTTCTGCGAAGGTGTTCATGATGATGTTGACATTAT 426
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 251 GATTGTAACATATCCAGGAGAGCTGCTTCTGCGAAGGTGTTCATGATGATGTTGACATTAT 310

Qy 427 ATTACTGCAAGATTAATGTGGTTACATATCTTATGTGACTGCCATTTTTGTTCTGGT 486
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 311 ATTACTGCAAGATTAATGTGGTTACATATCTTATGTGACTGCCATTTTTGTTCTGGT 370

Qy 487 AAACTGGAATATAAAGTGAAAGAACAAACATTTGAACATACTTAATGTATTTTATAGAA 546
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 371 AAACTGGAATATAAAGTGAAAGAACAAACATTTGAACATACTTAATGTATTTTATAGAA 430

Qy 547 CTTTGTAAACGAAAGGAGATTCATGTTTTAGAAGTCTGTCCTTTTTTATATCT 599
        |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 431 CTTTGTAAACGAAAGGAGATTCATGTTTTAGAAGTCTGTCCTTTTTTATATCT 483

RESULT 7
US-09-880-107-687/c
; Sequence 687, Application US/09880107
; Patent No. US20020142981A1
; GENERAL INFORMATION:
;   APPLICANT: Horne, Darci T.
;   APPLICANT: Vockley, Joseph G.
;   APPLICANT: Scherf, Uwe
;   APPLICANT: Gene Logic, Inc.
;   TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
;   FILE REFERENCE: 44921-5028-WO
;   CURRENT APPLICATION NUMBER: US/09/880,107
;   CURRENT FILING DATE: 2001-06-14
;   PRIOR APPLICATION NUMBER: US 60/211,379
;   PRIOR FILING DATE: 2000-06-14
;   PRIOR APPLICATION NUMBER: US 60/237,054
;   PRIOR FILING DATE: 2000-10-02
;   NUMBER OF SEQ ID NOS: 3950
;   SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 687
;   LENGTH: 406
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: Genbank Accession No. US20020142981A1 AA287347
US-09-880-107-687

Query Match   45.8%; Score 382.4; DB 9; Length 406;
Best Local Similarity   99.7%; Pred. No. 1.1e-83;
Matches   383; Conservative   0; Mismatches   1; Indels   0; Gaps   0;

Qy 452 CATATCTTTATGTACTGCCATTTTTGTTCCTGGTAAACTGGAATATAAAGTGAATATAAAGTGAAAGAAC 511
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 406 CATATCCTTATGTACTGCCATTTTTGTTCCTGGTAAACTGGAATATAAAGTGAATATAAAGTGAAAGAAC 347

Qy 512 AAACATTTGAACATACTTAATGTATTTTATAGAACTTTGTAAACGAAAGGAGATTCATG 571
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 346 AAACATTTGAACATACTTAATGTATTTTATAGAACTTTGTAAACGAAAGGAGATTCATG 287

Qy 572 TTTTAGAAGTCTGTCCTTTTTTATATCTTGAAAGAAAATCTATGTATGATGCTATAAAAT 631
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 286 TTTTAGAAGTCTGTCCTTTTTTATATCTTGAAAGAAAATCTATGTATGATGCTATAAAAT 227

Qy 632 AAATCCTATTATTTTCTCAGGAATCTGGTTAGGAATTGCAGGCAATGAGATTTTTTGCG 691
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 226 AAATCCTATTATTTTCTCAGGAATCTGGTTAGGAATTGCAGGCAATGAGATTTTTTGCG 167

Qy 692 GGGCAGGGATGGGAATGTTTGTTCATAAAATAATTAGACATTTCTATAGATATTTGACAT 751
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 166 GGGCAGGGATGGGAATGTTTGTTCATAAAATAATTAGACATTTCTATAGATATTTGACAT 107

Qy 752 TCTGCGAAAGCAACAAGCAAACTGAAGACAAACCCTATGAGAAATATATATGATGTTTAT 811
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 106 TCTGCGAAAGCAACAAGCAAACTGAAGACAAACCCTATGAGAAATATATATGATGTTTAT 47

Qy 812 GTAATAAAGACATGTAACTGTCTT 835
        ||||||||||||||||||||||||
Db   46 GTAATAAAGACATGTAACTGTCTT 23

RESULT 8
US-10-085-783A-21733
; Sequence 21733, Application US/10085783A
; Publication No. US20040037841A1
; GENERAL INFORMATION:
;   APPLICANT: ChondroGene Inc.
;   APPLICANT: Liew, C.C.
;   TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis
;   FILE REFERENCE: 4231/2002

CURRENT APPLICATION NUMBER: US/10/085,783A
CURRENT FILING DATE: 2002-02-28
PRIOR APPLICATION NUMBER: US 60/305,340
PRIOR FILING DATE: 2001-07-13
PRIOR APPLICATION NUMBER: US 60/275,017
PRIOR FILING DATE: 2001-03-12
PRIOR APPLICATION NUMBER: US 60/271,955
PRIOR FILING DATE: 2001-02-28
NUMBER OF SEQ ID NOS: 58994
SOFTWARE: PatentIn version 3.2
SEQ ID NO 21733
LENGTH: 400
TYPE: DNA
ORGANISM: Human
FEATURE:
NAME/KEY: misc_feature
LOCATION: (23)..(23)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc_feature
LOCATION: (48)..(48)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc_feature
LOCATION: (328)..(328)
OTHER INFORMATION: n is a, c, g, or t
-10-085-783A-21733

Query Match   45.6%;   Score 381; DB 12;   Length 400;
Best Local Similarity 96.8%;   Pred. No. 2.4e-83;
Matches 387; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

121 TGAAGAAGTAACCGGGTCCAGACCCACGCGGCGCCAGTTCTCCGGCGGGAAGGAAAAACC 180
    ||||||||||||||||||| | |||| ||||||||||||||||||||||||||||||
  1 TGAAGAAGTAACCGGGTCCATTNTCCACGCGGCGCCAGTTCTCCGGCNGGAAGGAAAAACC 60

181 GCGCAGAGAGGCAGCAATGAATGTGGATCACGAGGTTAACCTCTTAGTGGAGGAAATTCA 240
    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 61 GCGCAGAGAGGCAGCAATGAATGTGGATCACGAGGTTAACCTCTTAGTGGAGGAAATTCA 120

241 TCGTTTGGGTTCAAAAAATGCTGATGGAAAGTTAAGCCGTGAAATTTGGGGTCCTCTTCCG 300
    |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
121 TCGTTTGGGTTCAAAAAATGCTGATGGAAAGTTAAGCCGTGAAATTTGGGGTCCTCTTCCG 180

301 TGATGATAAATGTGCCAACCTCTTGAGCATTGGTAGGAACTCTTAAAGCTGCAAAACG 360
    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
181 TGATGATAAATGTGCCAACCTCTTGAGCATTGGTAGGAACTCTTAAAGCTGCAAAACG 240

361 AAGGAAGATTGTAACATATCCAGGAGAGCTGCTTCTGCAAGGTGTTCATGATGATGTTGA 420
    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
241 AAGGAAGATTGTAACATATCCAGGAGAGCTGCTTCTGCAAGGTGTTCATGATGATGTTGA 300

421 CATTATATTACTGCAAGATTAATGTGGTTTACATATCTTTATGTACTGCCATTTTTTGTT 480
    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
301 CATTATATTACTGCAAGATTAATGTGGNTTACATATCTTTATGTACTGCCATTTTTTGTT 360

481 TCTGTAAACTGGAATATAAAGTGAAAGAACAAACATTTG 520
    |||||||||||||||||||||||||||||||||||||||
361 TCTGTAAACTGGAATATAAAGTGAAAGAACAAACATTTG 400

RSULT 9
3-10-242-535A-21733
Sequence 21733, Application US/10242535A
Publication No. US20040013663A1
GENERAL INFORMATION:
APPLICANT: ChondroGene Inc.
APPLICANT: Liew, C.C.
TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis
FILE REFERENCE: 4231/2005
CURRENT APPLICATION NUMBER: US/10/242,535A
CURRENT FILING DATE: 2002-09-12
PRIOR APPLICATION NUMBER: US 10/085,783

PRIOR FILING DATE: 2002-02-28
PRIOR APPLICATION NUMBER: US 60/305,340
PRIOR FILING DATE: 2001-07-13
PRIOR APPLICATION NUMBER: US 60/275,017
PRIOR FILING DATE: 2001-03-12
PRIOR APPLICATION NUMBER: US 60/271,955
PRIOR FILING DATE: 2001-02-28
NUMBER OF SEQ ID NOS: 58994
SOFTWARE: PatentIn version 3.2
SEQ ID NO 21733
LENGTH: 400
TYPE: DNA
ORGANISM: Human
FEATURE:
NAME/KEY: misc_feature
LOCATION: (23)..(23)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc_feature
LOCATION: (48)..(48)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc_feature
LOCATION: (328)..(328)
OTHER INFORMATION: n is a, c, g, or t
US-10-242-535A-21733

Query Match   45.6%;   Score 381; DB 15;   Length 400;
Best Local Similarity 96.8%;   Pred. No. 2.4e-83;
Matches 387; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

Qy 121 TGAAGAAGTAACCGGGTCCAGACCCACGCGGCGCCAGTTCTCCGGCGGGAAGGAAAAACC 180
        ||||||||||||||||||| | |||| ||||||||||||||||||||||||||||||
Db   1 TGAAGAAGTAACCGGGTCCATTNTCCACGCGGCGCCAGTTCTCCGGCNGGAAGGAAAAACC 60

Qy 181 GCGCAGAGAGGCAGCAATGAATGTGGATCACGAGGTTAACCTCTTAGTGGAGGAAATTCA 240
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 61 GCGCAGAGAGGCAGCAATGAATGTGGATCACGAGGTTAACCTCTTAGTGGAGGAAATTCA 120

Qy 241 TCGTTTGGGTTCAAAAAATGCTGATGGAAAGTTAAGCCGTGAAATTTGGGGTCCTCTTCCG 300
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 121 TCGTTTGGGTTCAAAAAATGCTGATGGAAAGTTAAGCCGTGAAATTTGGGGTCCTCTTCCG 180

Qy 301 TGATGATAAATGTGCCAACCTCTTGAAGCATTGGTAGGAACTCTTAAAGCTGCAAAACG 360
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 181 TGATGATAAATGTGCCAACCTCTTGAAGCATTGGTAGGAACTCTTAAAGCTGCAAAACG 240

Qy 361 AAGGAAGATTGTAACATATCCAGGAGAGCTGCTTCTGCAAGGTGTTCATGATGATGTTGA 420
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 241 AAGGAAGATTGTAACATATCCAGGAGAGCTGCTTCTGCAAGGTGTTCATGATGATGTTGA 300

Qy 421 CATTATATTACTGCAAGATTAATGTGGTTTACATATCTTTATGTACTGCCATTTTTTGTT 480
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 301 CATTATATTACTGCAAGATTAATGTGGNTTACATATCTTTATGTACTGCCATTTTTTGTT 360

Qy 481 TCTGTAAACTGGAATATAAAGTGAAAGAACAAACATTTG 520
        |||||||||||||||||||||||||||||||||||||||
Db 361 TCTGTAAACTGGAATATAAAGTGAAAGAACAAACATTTG 400

RESULT 10
US-09-878-178-2101/c
Sequence 2101, Application US/09878178
Patent No. US2002017755Aa1
GENERAL INFORMATION:
APPLICANT: Jiang, Yuqiu
APPLICANT: Harlocker, Susan L.
APPLICANT: Secrist, Heather
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
TITLE OF INVENTION: AND DIAGNOSIS OF COLON CANCER
FILE REFERENCE: 210121.527
CURRENT APPLICATION NUMBER: US/09/878,178
CURRENT FILING DATE: 2001-06-08

NUMBER OF SEQ ID NOS: 2237
SOFTWARE: FastSEQ for Windows Version 4.0
SEQ ID NO 2101
  LENGTH: 365
  TYPE: DNA
  ORGANISM: Homo sapien
  FEATURE:
  NAME/KEY: misc_feature
  LOCATION: (1)...(365)
  OTHER INFORMATION: n = A,T,C or G
-09-878-178-2101

Query Match   43.6%;   Score 364; DB 9;   Length 365;
Best Local Similarity 99.7%;   Pred. No. 3.3e-79;
Matches 364; Conservative 0; Mismatches 1; Indels 1; Gaps 0;

```
Qy  465 ACTGCCATTTTTGTTCTCGGTAAACTGGAATATGAAGTGAAAGAACAAACATTTGAACA 524
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  365 ACTGCCATTTTTGTTCTCGGTAAACTGGAATATAAAGTGAAAGAACAAACATTTGAACA 306

Qy  525 TACTTAATGTATTTTATAGAACTTGTAAACGAAAGGAGATTCATGTTTTAGAAGTCTG 584
        |||||||||||||||||||||||||||||| |||||||||||||||||||||||||||
Db  305 TACTTAATGTATTTTATAGAACTTGTAAACNAAAGGAGATTCATGTTTTAGAAGTCTG 246

Qy  585 TCCTTTTTTATATCTTGAAAGAAAATCTATGTGATGCTATAAAATAAATCCTATTATT 644
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  245 TCCTTTTTTATATCTTGAAAGAAAATCTATGTGATGCTATAAAATAAATCCTATTATT 186

Qy  645 TTTCTCAGGAATCTGGTTAGGAATTGCAGGCAATGAGATTTTTGCGGGGCAGGGATGGG 704
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  185 TTTCTCAGGAATCTGGTTAGGAATTGCAGGCAATGAGATTTTTGCGGGGCAGGGATGGG 126

Qy  705 AATGTTTGTTCATAAATAATTAGACATTTCTATAGATATTTGACATTCTGCGAAAGCAA 764
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  125 AATGTTTGTTCATAAATAATTAGACATTTCTATAGATATTTGACATTCTGCGAAAGCAA 66

Qy  765 CAAGCAAACTGAAGACCAACTCCTATGAGAAATATTATGATGTTTATGTAATAAAGACAT 824
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   65 CAAGCAAACTGAAGACCAACTCCTATGAGAAATATTATGATGTTTATGTAATAAAGACAT 6

Qy  825 GTAAC 829
        |||||
Db    5 GTAAC 1
```

ESULT 11
S-10-046-935-2101/c
Sequence 2101, Application US/10046935
Publication No. US20020156011A1
GENERAL INFORMATION:
  APPLICANT: Jiang, Yuqiu
  APPLICANT: Harlocker, Susan L.
  APPLICANT: Secrist, Heather
  APPLICANT: Wang, Aijun
  APPLICANT: Stolk, John A.
  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
  TITLE OF INVENTION: AND DIAGNOSIS OF COLON CANCER
  FILE REFERENCE: 210121.527C1
  CURRENT APPLICATION NUMBER: US/10/046,935
  CURRENT FILING DATE: 2002-01-15
  NUMBER OF SEQ ID NOS: 2239
  SOFTWARE: FastSEQ for Windows Version 4.0
SEQ ID NO 2101
  LENGTH: 365
  TYPE: DNA
  ORGANISM: Homo sapiens
  FEATURE:
  NAME/KEY: misc_feature
  LOCATION: 273
  OTHER INFORMATION: n = A,T,C or G
US-10-046-935-2101

Query Match   43.6%;   Score 364;   DB 13;   Length 365;

Best Local Similarity 99.7%;   Pred. No. 3.3e-79;
Matches 364; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy  465 ACTGCCATTTTTGTTCTCGGTAAACTGGAATATGAAGTGAAAGAACAAACATTTGAACA 524
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  365 ACTGCCATTTTTGTTCTCGGTAAACTGGAATATAAAGTGAAAGAACAAACATTTGAACA 306

Qy  525 TACTTAATGTATTTTATAGAACTTGTAAACGAAAGGAGATTCATGTTTTAGAAGTCTG 584
        |||||||||||||||||||||||||||||| |||||||||||||||||||||||||||
Db  305 TACTTAATGTATTTTATAGAACTTGTAAACNAAAGGAGATTCATGTTTTAGAAGTCTG 246

Qy  585 TCCTTTTTTATATCTTGAAAGAAAATCTATGTGATGCTATAAAATAAATCCTATTATT 644
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  245 TCCTTTTTTATATCTTGAAAGAAAATCTATGTGATGCTATAAAATAAATCCTATTATT 186

Qy  645 TTTCTCAGGAATCTGGTTAGGAATTGCAGGCAATGAGATTTTTGCGGGGCAGGGATGGG 704
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  185 TTTCTCAGGAATCTGGTTAGGAATTGCAGGCAATGAGATTTTTGCGGGGCAGGGATGGG 126

Qy  705 AATGTTTGTTCATAAATAATTAGACATTTCTATAGATATTTGACATTCTGCGAAAGCAA 764
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  125 AATGTTTGTTCATAAATAATTAGACATTTCTATAGATATTTGACATTCTGCGAAAGCAA 66

Qy  765 CAAGCAAACTGAAGACCAACTCCTATGAGAAATATTATGATGTTTATGTAATAAAGACAT 824
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   65 CAAGCAAACTGAAGACCAACTCCTATGAGAAATATTATGATGTTTATGTAATAAAGACAT 6

Qy  825 GTAAC 829
        |||||
Db    5 GTAAC 1
```

RESULT 12
US-10-146-502-2101/c
Sequence 2101, Application US/10146502
Publication No. US20030069180A1
GENERAL INFORMATION:
  APPLICANT: Jiang, Yuqiu
  APPLICANT: Harlocker, Susan L.
  APPLICANT: Secrist, Heather
  APPLICANT: Wang, Aijun
  APPLICANT: Stolk, John A.
  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
  TITLE OF INVENTION: AND DIAGNOSIS OF COLON CANCER
  FILE REFERENCE: 210121.527C2
  CURRENT APPLICATION NUMBER: US/10/146,502
  CURRENT FILING DATE: 2002-05-14
  NUMBER OF SEQ ID NOS: 2241
  SOFTWARE: FastSEQ for Windows Version 4.0
SEQ ID NO 2101
  LENGTH: 365
  TYPE: DNA
  ORGANISM: Homo sapiens
  FEATURE:
  NAME/KEY: misc_feature
  LOCATION: 273
  OTHER INFORMATION: n = A,T,C or G
US-10-146-502-2101

Query Match   43.6%;   Score 364;   DB 14;   Length 365;
Best Local Similarity 99.7%;   Pred. No. 3.3e-79;
Matches 364; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy  465 ACTGCCATTTTTGTTCTCGGTAAACTGGAATATGAAGTGAAAGAACAAACATTTGAACA 524
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  365 ACTGCCATTTTTGTTCTCGGTAAACTGGAATATAAAGTGAAAGAACAAACATTTGAACA 306

Qy  525 TACTTAATGTATTTTATAGAACTTGTAAACGAAAGGAGATTCATGTTTTAGAAGTCTG 584
        |||||||||||||||||||||||||||||| |||||||||||||||||||||||||||
Db  305 TACTTAATGTATTTTATAGAACTTGTAAACNAAAGGAGATTCATGTTTTAGAAGTCTG 246

Qy  585 TCCTTTTTTATATCTTGAAAGAAAATCTATGTGATGCTATAAAATAAATCCTATTATT 644
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

245 TCCTTTTTTATATCTTGAAAGAAAATCTATGTATGATGCTATAAAAATAAATCCTATTATT 186
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

645 TTTCTCGGAATCTGTTAGGAATTGCAGGCAATGAGATTTTTGCGGGGCAGGGGATGGG 704
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
185 TTTCTCAGGAATCTGTTAGGAATTGCAGGCAATGAGATTTTTGCGGGGCAGGGGATGGG 126

705 AATGTTTGTTCATAAATAATTAGACATTTTCTATAGATATTTGACATTCTCGGGAAGCAA 764
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
125 AATGTTTGTTCATAAATAATTAGACATTTTCTATAGATATTTGACATTCTCGGAAAGCAA 66

765 CAAGCAAACTGAAGACCAACCTCCTATGAGAAAATATTATGATGTTTATGTAATAAAGACAT 824
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
65 CAAGCAAACTGAAGACCAACCTCCTATGAGAAATATTTATGATGTTTATGTAATAAAGACAT 6

825 GTAAC 829
        |||||
5 GTAAC 1

SULT 13
-09-918-995-26075
Sequence 26075, Application US/09918995
Publication No. US20030073623A1
GENERAL INFORMATION:
APPLICANT: Hyseq, Inc.
TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
TITLE OF INVENTION: FROM VARIOUS CDNA LIBRARIES
FILE REFERENCE: 20411-756
CURRENT APPLICATION NUMBER: US/09/918,995
CURRENT FILING DATE: 2001-07-30
PRIOR APPLICATION NUMBER: US/09/235,076
PRIOR FILING DATE: 1999-01-20
NUMBER OF SEQ ID NOS: 38054
SOFTWARE: FastSEQ for Windows Version 3.0
SEQ ID NO 26075
    LENGTH: 458
    TYPE: DNA
    ORGANISM: Homo sapiens
    FEATURE:
    NAME/KEY: misc_feature
    LOCATION: (1)...(458)
    OTHER INFORMATION: n = A,T,C or G
3-09-918-995-26075

Query Match   28.4%; Score 237; DB 10; Length 458;
Best Local Similarity 100.0%; Pred. No. 5.5e-48;
Matches 237; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

188 GAGGCAGCAATGAATGTGGATCACGAGGTTAACCTCTTAGTGAGGAAATTCATCGTTTG 247
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
222 GAGGCAGCAATGAATGTGGATCACGAGGTTAACCTCTTAGTGGGAGGAAATTCATCGTTTG 281

248 GGTTCAAAAAATGCTGATGGAAAGTTAAGCGTGAAATTGGGCGTCCTCCTTCCGTGATGAT 307
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
282 GGTTCAAAAAATGCTGATGGAAAGTTAAGCGTGAAATTTGGGGTCCTCCTTCCGTGATGAT 341

308 AAATGTGCCAACCTCTTTGAGGCAATGGTAGGAACTCCTAAAAGCTGCAAAACGAAGGAAG 367
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
342 AAATGTGCCAACCTCTTTGAGGCAATGGTAGGAACTCCTAAAAGCTGCAAAACGAAGGAAG 401

368 ATTGTAACATATCCAGGAGAGCTGCTTCTGCAAGGTGTTCATGATGATGTTGACATT 424
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
402 ATTGTAACATATCCAGGAGAGCTGCTTCTGCAAGGTGTTCATGATGATGTTGACATT 458

ESULT 14
S-09-925-299-318
Sequence 318, Application US/09925299
Patent No. US20020055627A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
FILE REFERENCE: PA102

CURRENT APPLICATION NUMBER: US/09/925,299
CURRENT FILING DATE: 2001-08-10
PRIOR APPLICATION NUMBER: PCT/US00/05883
PRIOR FILING DATE: 2000-03-08
PRIOR APPLICATION NUMBER: 60/124,270
PRIOR FILING DATE: 1999-03-12
NUMBER OF SEQ ID NOS: 1556
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 318
    LENGTH: 388
    TYPE: DNA
    ORGANISM: Homo sapiens
    FEATURE:
    NAME/KEY: misc_feature
    LOCATION: (1)
    OTHER INFORMATION: n equals a,t,g, or c
    NAME/KEY: misc_feature
    LOCATION: (2)
    OTHER INFORMATION: n equals a,t,g, or c
    NAME/KEY: misc_feature
    LOCATION: (3)
    OTHER INFORMATION: n equals a,t,g, or c
    NAME/KEY: misc_feature
    LOCATION: (14)
    OTHER INFORMATION: n equals a,t,g, or c
    NAME/KEY: misc_feature
    LOCATION: (143)
    OTHER INFORMATION: n equals a,t,g, or c
    NAME/KEY: misc_feature
    LOCATION: (144)
    OTHER INFORMATION: n equals a,t,g, or c
    NAME/KEY: misc_feature
    LOCATION: (146)
    OTHER INFORMATION: n equals a,t,g, or c
    NAME/KEY: misc_feature
    LOCATION: (150)
    OTHER INFORMATION: n equals a,t,g, or c
    NAME/KEY: misc_feature
    LOCATION: (154)
    OTHER INFORMATION: n equals a,t,g, or c
    NAME/KEY: misc_feature
    LOCATION: (159)
    OTHER INFORMATION: n equals a,t,g, or c
    NAME/KEY: misc_feature
    LOCATION: (160)
    OTHER INFORMATION: n equals a,t,g, or c
    NAME/KEY: misc_feature
    LOCATION: (163)
    OTHER INFORMATION: n equals a,t,g, or c
    NAME/KEY: misc_feature
    LOCATION: (168)
    OTHER INFORMATION: n equals a,t,g, or c
    NAME/KEY: misc_feature
    LOCATION: (171)
    OTHER INFORMATION: n equals a,t,g, or c
    NAME/KEY: misc_feature
    LOCATION: (174)
    OTHER INFORMATION: n equals a,t,g, or c
    NAME/KEY: misc_feature
    LOCATION: (196)
    OTHER INFORMATION: n equals a,t,g, or c
    NAME/KEY: misc_feature
    LOCATION: (198)
    OTHER INFORMATION: n equals a,t,g, or c
    NAME/KEY: misc_feature
    LOCATION: (223)
    OTHER INFORMATION: n equals a,t,g, or c
    NAME/KEY: misc_feature
    LOCATION: (226)
    OTHER INFORMATION: n equals a,t,g, or c
    NAME/KEY: misc_feature
    LOCATION: (227)
    OTHER INFORMATION: n equals a,t,g, or c

FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc_feature
LOCATION: (2)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc_feature
LOCATION: (3)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc_feature
LOCATION: (14)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc_feature
LOCATION: (143)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc_feature
LOCATION: (144)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc_feature
LOCATION: (146)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc_feature
LOCATION: (150)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc_feature
LOCATION: (154)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc_feature
LOCATION: (159)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc_feature
LOCATION: (160)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc_feature
LOCATION: (163)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc_feature
LOCATION: (168)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc_feature
LOCATION: (171)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc_feature
LOCATION: (174)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc_feature
LOCATION: (196)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc_feature
LOCATION: (198)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc_feature
LOCATION: (223)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc_feature
LOCATION: (226)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc_feature
LOCATION: (227)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc_feature
LOCATION: (234)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc_feature
LOCATION: (240)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc_feature
LOCATION: (242)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc_feature
LOCATION: (255)
OTHER INFORMATION: n equals a,t,g, or c

NAME/KEY: misc_feature
LOCATION: (234)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc_feature
LOCATION: (240)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc_feature
LOCATION: (242)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc_feature
LOCATION: (255)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc_feature
LOCATION: (280)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc_feature
LOCATION: (314)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc_feature
LOCATION: (359)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc_feature
LOCATION: (366)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc_feature
LOCATION: (368)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc_feature
LOCATION: (377)
OTHER INFORMATION: n equals a,t,g, or c

S-09-925-299-318

Query Match   22.6%;   Score 188.4; DB 9;   Length 388;
Best Local Similarity 82.8%; Pred. No. 4.2e-36;
Matches 198; Conservative 0; Mismatches 41; Indels 0; Gaps 0;

```
Qy    134 CGGGTCCAGACCCACGCGCGCCAGTTCTCCGGCGGAAGAACCGGCAGAGAGGCA 193
          ||||||||||||||||||||||||| || ||  || |||||||||  | |||||
Db    138 CCGGTNNANACCNATGNGGGGNNAGNTCTCNGGNGGNAAGGAAAATCGTGCAAAGAGGNA 197

Qy    194 GCAATGAAGTGTGGATCACGAGGTTAACCTCTTAGTGAGGAAATTCATCGTTGGGTTCA 253
          ||||||  | ||||||| ||||| || ||  |||| ||| ||| |||| ||||||||
Db    198 NTAATGAATGTGGATCACGAGGATANCCNNTTAGTGNAGGAANTNCATCGTTGGGTNTA 257

Qy    254 AAAAATGTGATGGAAAGTTAAGCGTGAAAATTTGGGGTCCTTCCGTGATGATAAATGT 313
          ||||||||||||||||||||||||| |||||||||||||||||| ||||||||||
Db    258 AAAAATGCTGATGGAAAGTTAANTGTGAAAATTGGGGTCCTCTTTCGTGATGATAANTGA 317

Qy    314 GCCAACCTCTTGAAGCATTGGTAGGAACTCTTAAAGCTGCAAAACGAAGGAAGATTGT 372
          ||||||||||||||||||||||||||||||||||||| ||||| ||||||||||||
Db    318 GCCAACCTCTTTGAAGCATTGGTAGGAACTCTTAAAGCTGTNAAACGANANAAGAATGT 376
```

NAME/KEY: misc_feature
LOCATION: (280)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc_feature
LOCATION: (314)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc_feature
LOCATION: (359)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc_feature
LOCATION: (366)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc_feature
LOCATION: (368)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc_feature
LOCATION: (377)
OTHER INFORMATION: n equals a,t,g, or c
-09-925-299-318

Query Match   22.6%;   Score 188.4; DB 10; Length 388;
Best Local Similarity   82.8%;   Pred. No. 4.2e-36;
Matches 198; Conservative   0; Mismatches 41; Indels   0; Gaps   0;

```
    134 CGGGTCCAGACCCACGCGGCGCCAGTTCTCCGGCGGGGAAGGAAAAACCGCGCAGAGAGGCA 193
        | |||  |  ||||   |||   |||||||   ||||||| ||||||| ||||||||
    138 CCGGTNNANACCNATGNGGGGNNAGNTCTCNGGNGNGNAAGGAAAAATCGTGCAAAGAGGNA 197

    194 GCAATGAATGTGGATCACGAGGTTAACCTCTTAGTGGAGGAAATTCATCGTTTGGGTTCA 253
        |||||||||||||||||||   |||   |||||   ||||   ||||||||||| |
    198 NTAATGAATGTGGATCACGAGGATANCCNNTTAGTGNAGGAANTNCATCGTTTGGGTNTA 257

    254 AAAAATGCTGATGAAAGTTAAGCGTGAAATTTGGGGTCCTCTTCCGTGATGATAAATGT 313
        |||||||||||||||| |||||||||||||||||||||||||||||||||| |
    258 AAAAATGCTGATGGAAAGTTAANTGTGAAAAATTGGGGTCCTCTTTCGTGATGATAANTGA 317

    314 GCCAACCTCTTGAAGCATTGGTAGGAACTCTTAAAGCTGCAAAACGAAGGAAGATTGT 372
        ||||||||||||||||||||||||||||||||||| |  ||||||||||||||| | | |
    318 GCCAACCTCTTTGAAGCATTGGTAGGAACTCTTAAAGCTGTNAAACGANANAAGAATGT 376
```

arch completed: March 17, 2004, 14:31:49
b time : 385 secs

```
            GenCore version 5.1.6
    Copyright (c) 1993 - 2004  Compugen Ltd.

protein - nucleic search, using frame_plus_p2n model

n on:    March 17, 2004, 16:42:09 ; Search time 1964 Seconds
                                (without alignments)
                       1231.587 Million cell updates/sec

tle:     US-09-648-310-4
rfect score:  410
quence:  1 MNVDHEVNLLVEEIHRLGSK.........PGELLLQGVHDDVDILLQD 81

oring table:  BLOSUM62
         Xgapop  10.0 ,  Xgapext  0.5
         Ygapop  10.0 ,  Ygapext  0.5
         Fgapop   6.0 ,  Fgapext  7.0
         Delop    6.0 ,  Delext   7.0

arched:  27513289 seqs, 1493109276 residues

tal number of hits satisfying chosen parameters:    55026578

nimum DB seq length: 0
ximum DB seq length: 2000000000

st-processing: Minimum Match 0%
               Maximum Match 100%
               Listing first 45 summaries

mmand line parameters:
ODEL=frame+p2n.model -DEV=xlh
=/cgn2_1/USPTO_spool/US09648310/runat_15032004_101254_16573/app_query.fasta_1.263
B=EST -QFMT=fastap -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
NITs=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
OCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
UTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
SER=US09648310 @CGN_1_1_3437 @runat_15032004_101254_16573 -NCPU=6 -ICPU=3
O_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
EV_TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
GAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

tabase :   EST:*
      1:  em_estba:*
      2:  em_esthum:*
      3:  em_estin:*
      4:  em_estmu:*
      5:  em_estov:*
      6:  em_estpl:*
      7:  em_estro:*
      8:  em_htc:*
      9:  gb_esti:*
     10:  gb_est2:*
     11:  gb_htc:*
     12:  gb_est3:*
     13:  gb_est4:*
     14:  gb_est5:*
     15:  em_estom:*
     16:  em_estom:*
     17:  em_gss_hum:*
     18:  em_gss_inv:*
     19:  em_gss_pln:*
     20:  em_gss_vrt:*
     21:  em_gss_fun:*
     22:  em_gss_mam:*
     23:  em_gss_mus:*
     24:  em_gss_pro:*
     25:  em_gss_rod:*
     26:  em_gss_phg:*
     27:  em_gss_vrl:*
     28:  gb_gss1:*
     29:  gb_gss2:*
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result NO.	Score	Query Length	DB	% Match	ID	Description
1	410	477	12	100.0	BM752941	BM752941 K-EST0029
2	410	480	13	100.0	BU199007	BU199007 DCBCYC06
3	410	532	14	100.0	N39717	N39717 yx92d07.r1
4	410	542	12	100.0	BG944932	BG944932 ax56e07.x
5	410	544	13	100.0	BX474507	BX474507 DKFZp686E
6	410	567	13	100.0	BU198216	BU198216 DCBCNC04
7	410	571	10	100.0	BE145690	BE145690 ILO-HT020
8	410	576	14	100.0	CA335746	CA335746 NISC_lu03
9	410	599	9	100.0	AV753954	AV753954 AV753954
10	410	599	14	100.0	CB162901	CB162901 K-EST0223
11	410	601	9	100.0	AV714738	AV714738 AV714738
12	410	601	9	100.0	AV716706	AV716706 AV716706
13	410	604	13	100.0	BU581717	BU581717 in40c11.y
14	410	634	9	100.0	AI133406	AI133406 HA2013 Hu
15	410	641	12	100.0	BG563686	BG563686 60258452
16	410	648	13	100.0	BQ777108	BQ777108 i143e02.y
17	410	666	12	100.0	BG745050	BG745050 602723144
18	410	676	13	100.0	BX110370	BX110370 BX110370
19	410	684	12	100.0	BG744405	BG744405 602723144
20	410	694	12	100.0	BG500857	BG500857 602547312
21	410	698	10	100.0	BE891760	BE891760 601434485
22	410	702	9	100.0	AV716325	AV716325 AV716325
23	410	717	13	100.0	BU598760	BU598760 AGENCOURT
24	410	741	9	100.0	AV717724	AV717724 AV717724
25	410	752	9	100.0	AV716363	AV716363 AV716363
26	410	755	12	100.0	BG530613	BG530613 60250008
27	410	763	9	100.0	AV756659	AV756659 AV756659
28	410	765	13	100.0	BU953438	BU953438 AGENCOURT
29	410	791	12	100.0	BG528919	BG528919 602579350
30	410	797	12	100.0	BI859694	BI859694 60338683
31	410	812	14	100.0	CB958631	CB958631 60338283
32	410	816	13	100.0	BU849360	BU849360 AGENCOURT
33	410	845	11	100.0	AF161398	AF161398 Homo sapi
34	410	876	11	100.0	AF116682	AF116682 Homo sapi
35	410	884	12	100.0	BG164526	BG164526 602342067
36	410	906	14	100.0	CD521898	CD521898 AGENCOURT
37	410	1095	12	100.0	BG114255	BG114255 602285822
38	410	1126	12	100.0	BM542613	BM542613 AGENCOURT
39	409	538	12	99.8	BG385103	BG385103 30670G MA
40	409	586	13	99.8	BX670691	BX670691 BX670691
41	409	647	14	99.8	CB477970	CB477970 jns28_D03
42	409	716	13	99.3	BX677255	BX677255 BX677255
43	407	530	14	99.3	CF340773	CF340773 TgESTzyj4
44	406	492	9	99.0	AV712262	AV712262 AV712262
45	404	674	9	98.5	AW027067	AW027067 wt71f11.x

ALIGNMENTS

RESULT 1
BM752941
LOCUS   BM752941   477 bp   mRNA   linear   EST 04-MAR-2002
DEFINITION K-EST0029681 S3SNU16 Homo sapiens cDNA clone S3SNU16-30-F01 5',
            mRNA sequence.
ACCESSION   BM752941
VERSION   BM752941.1 GI:19082559
KEYWORDS   EST.
SOURCE   Homo sapiens (human)
  ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 477)

RESULT 2
BU199007
LOCUS   BU199007   480 bp   mRNA   linear   EST 05-SEP-2002
DEFINITION   DCBCYC06 DCB Homo sapiens cDNA, mRNA sequence.
ACCESSION   BU199007
VERSION   BU199007.1 GI:22717475
KEYWORDS   EST.
SOURCE   Homo sapiens (human)
ORGANISM   Homo sapiens
                Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 480)
AUTHORS   Xu,X., Gu,J., Liu,F., Qu,J., Zhao,M., Li,Y., Huang,Q., Zhou,J.,
                Song,H., Gu,Y., Yang,Y., Gao,G., Xiao,H., Li,N., Qian,B., Gao,X.,
                Cheng,Z., Xu,S., Gu,W., Tu,Y., Jia,J., Fu,G., Ren,S., Zhong,M.,
                Lu,G., Cheng,Z. and Han,Z.
TITLE   Homo sapiens cDNA DCB clones
JOURNAL   Unpublished (2000)
COMMENT   Contact: Zeguang Han
                Chinese National Human Genome Center at Shanghai
                351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
                201203, P. R. China
                Tel: 86-21-50801919(ex.45)
                Fax: 86-21-50801922
                Email: hanzg@chgc.sh.cn.
FEATURES   Location/Qualifiers
     source   1..480
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /cell_type="dendritic cells"
                     /dev_stage="mature"
                     /lab_host="BM25.8"
                     /clone_lib="DCB"
                     /note="Vector: pTriplEx2; Site_1: sfiIA; Site_2: sfiIB"

ORIGIN

Alignment Scores:
Pred. No.:   2.58e-45   Length: 480
Score: 410.00   Matches: 81
Percent Similarity: 100.00%   Conservative: 0
Best Local Similarity: 100.00%   Mismatches: 0
Query Match: 100.00%   Indels: 0
DB: Gaps: 0

US-09-648-310-4 (1-81) x BU199007 (1-480)

Qy   1 MetAsnValAspHisGluValAsnLeuValGluGluIleHisArgLeuGlySerLys 20
Db   189 ATGAATGTGGATCACGAGGTTAACCTCTTAGTGGAGGAAATTCATCGTTTGGGTTCAAAA 248

Qy   21 AsnAlaAspGlyLysLeuSerValLysPheGlyValLeuPheArgAspLysCysAla 40
Db   249 AATTGCTGATGGAAAGTTAAGCGTGAAAATTTGGGGTCCTCTTCCGTGATGATAAATGTGCC 308

Qy   41 AsnLeuPheGluAlaLeuValGlyThrLeuLysAlaAlaLysArgArgLysIleValThr 60
Db   309 AACCTCTTTGAAGCATTGGTAGGAACTCTTAAAGTGCAAAACGAGGAAGATTGTAACA 368

Qy   61 TyrProGlyLeuLeuLeuGlnGlyValHisAspAspValAspIleIleLeuLeuGln 80
Db   369 TATCCAGGAGAGCTGCTTCTGCAAGGTGTTCATGATGATGTTGACATTATTATTACTGCAA 428

Qy   81 Asp 81
Db   429 GAT 431

RESULT 3
N39717
LOCUS   N39717   532 bp   mRNA   linear   EST 22-JAN-1996
DEFINITION   yx92d07.r1 Soares melanocyte 2NbHM Homo sapiens cDNA clone
                IMAGE:269197 5', mRNA sequence.

TITLE   21C Frontier Korean EST Project 2001
JOURNAL   Unpublished (2002)
COMMENT   Contact: Kim YS
                Genome Research Center
                Korea Research Institute of Bioscience & Biotechnology
                52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
                Tel: +82-42-860-4470
                Fax: +82-42-860-4409
                Email: yongsung@mail.kribb.re.kr
                Plate: 30   row: F   column: 01
                High quality sequence stop: 477.
FEATURES   Location/Qualifiers
     source   1..477
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="S33NU16-30-F01"
                     /sex="F"
                     /tissue_type="Ascites"
                     /cell_type="Lymphoblast-like"
                     /cell_line="SNU-16"
                     /lab_host="Top10F'"
                     /clone_lib="S33SNU16"
                     /note="Organ: Stomach; Vector: pTZ18RP1; Site_1: EcoRI;
                     Site_2: NotI; The poly (A)+ RNA was dephosphorylated with
                     bacterial alkaline pyrophosphatase (BAP) and then decapped
                     with tabacco acid pyrophosphatase (TAP). The decapped
                     intact mRNA was ligated with DNA-RNA linker including EcoR
                     I site by treatment of T4 RNA ligase and the first strand
                     cDNA was synthesized from oligo dT-selected mRNA by
                     priming with dT-tailed vector. The dT-tailed vector was
                     adjusted to have about 60nt. The cDNA vector was
                     circularized with E. coli DNA ligase after digestion of
                     EcoRI which site is also included in vector. An RNA strand
                     converted to a DNA strand by Okayama-Berg method. The
                     obtained cDNA vectors were used for transformation method.
                     competent cells E. coli Top10F' by electroporation method.
                     The cDNA libraries constructed by this method are
                     full-length enriched cDNA library."

ORIGIN

Alignment Scores:
Pred. No.:   2.55e-45   Length: 477
Score: 410.00   Matches: 81
Percent Similarity: 100.00%   Conservative: 0
Best Local Similarity: 100.00%   Mismatches: 0
Query Match: 100.00%   Indels: 0
DB: Gaps: 0

JS-09-648-310-4 (1-81) x BM752941 (1-477)

Qy   1 MetAsnValAspHisGluValAsnLeuLeuValGluGluIleHisArgLeuGlySerLys 20
Db   139 ATGAATGTGGATCACGAGGTTAACCTCTTAGTGGAGGAAATTCATCGTTTGGGTTCAAAA 198

Qy   21 AsnAlaAspGlyLysLeuSerValLysPheGlyValLeuPheArgAspLysCysAla 40
Db   199 AATTGCTGATGGAAAGTTAAGCGTGAAAATTTGGGGTCCTCTTCCGTGATGATAAATGTGCC 258

Qy   41 AsnLeuPheGluAlaLeuValGlyThrLeuLysAlaAlaLysArgArgLysIleValThr 60
Db   259 AACCTCTTTGAAGCATTGGTAGGAACTCTTAAAGTGCAAAACGAGGAAGATTGTAACA 318

Qy   61 TyrProGlyLeuLeuLeuGlnGlyValHisAspAspValAspIleIleLeuLeuGln 80
Db   319 TATCCAGGAGAGCTGCTTCTGCAAGGTGTTCATGATGATGTTGACATTATTATTACTGCAA 378

Qy   81 Asp 81
Db   379 GAT 381

```
Db      270 TATCCAGGAGGAGCTGCTTCTGCAAGGTGTTCATGATGATGTTGACATTATTACTGCAA 329
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy       81 Asp 81
            |||
Db      330 GAT 332
```

RESULT 4
BG944932
LOCUS   BG944932   542 bp   mRNA   linear   EST 15-JAN-2003
DEFINITION ax56e07.x1 Hembase; Erythroid Progenitor Cells (LCB:ax library)
            Homo sapiens cDNA clone ax56e07 random, mRNA sequence.
ACCESSION   BG944932
VERSION   BG944932.1 GI:14344293
KEYWORDS   EST.
SOURCE   Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 542)
AUTHORS   Gubin,A.N., Njoroge,J.M., Bouffard,G.G. and Miller,J.L.
TITLE   Gene expression in proliferating human erythroid cells
JOURNAL   Genomics 59 (2), 168-177 (1999)
MEDLINE   99339981
PUBMED   10409428
COMMENT   Contact: Jeffery L. Miller
            Laboratory of Chemical Biology
            National Institute of Diabetes and Digestive and Kidney Diseases
            Building 10, Room 9B17, National Institutes of Health
            20892, USA
            Tel: 301 402 2373
            Fax: 301 435 5148
            Email: jm7f@nih.gov
            DNA Sequencing and analyses by National Institutes of Health
            Intramural Sequencing Center (NISC).
            Plate: 56 row: e column: 07
            Seq primer: -21M13 forward primer (ABI).
FEATURES   Location/Qualifiers
     source   1..542
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="ax56e07"
                     /sex="unknown"
                     /tissue_type="blood"
                     /cell_type="Erythroid Cells"
                     /cell_lines="Primary Culture of Peripheral Blood
                     Mononuclear Cells"
                     /dev_stage="Progenitor; EPO responsive CD71+++"
                     /lab_host="SOLR"
                     /clone_lib="Hembase; Erythroid Progenitor Cells (LCB:ax
                     library)"
                     /note="Organ: blood; Vector: Lambda ZAP II; Site 1: EcoRI;
                     Site 2: EcoRI; 65,000 proliferating erythroid cells from
                     the buffy coat of a blood donation were obtained by flow
                     cytometric separation after a 5-day culture period in the
                     presence of erythropoietin. Total RNA was purified from
                     the sorted cell population using TRIzol reagent. RNA (0.3
                     ug) was converted into double stranded cDNA using
                     Clontech's CapFinder cDNA Library Construction Kit
                     (Clontech) according to the manufacturer's protocol and
                     cloned into EcoRI digested Lambda Zap II vector
                     (Stratagene). The phage library was amplified once prior
                     to in vivo excision in SOLR cells. Individual colonies
                     were grown, and the cDNA inserts were sequenced in high
                     throughput (NIH intramural sequencing center
                     http://www.nisc.nih.gov/)."
ORIGIN

Alignment Scores:
Pred. No.:   3.03e-45   Length: 542
Score: 410.00   Matches: 81
Percent Similarity: 100.00%   Conservative: 0
```

---

ACCESSION   N39717
VERSION     N39717.1  GI:1163262
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 532)
AUTHORS     Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
            Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
            Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,
            Trevaskis,E., Waterston,R., Williamson,A., Wohldmann,P. and
            Wilson,R.
TITLE       The WashU-Merck EST Project
JOURNAL     Unpublished (1995)
COMMENT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            High quality sequence stops: 414
            Source: IMAGE Consortium, LLNL
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Seq primer: T7
            High quality sequence stop: 414.
FEATURES             Location/Qualifiers
     source          1..532
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="GDB:3878839"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:269197"
                     /sex="Male"
                     /tissue_type="melanocyte"
                     /lab_host="DH10B (ampicillin resistant)"
                     /clone_lib="Soares melanocyte 2NbHM"
                     /note="Vector: pT7T3D (Pharmacia) with a modified
                     polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA
                     was primed with a Not I - oligo(dT) primer [5'
                     TGTTACCAATTCGAAGTGGGAGCGGCCGCAGTTTTTTTTTTTTTTTTT 3'],
                     double-stranded cDNA was size selected, ligated to Eco RI
                     adapters (Pharmacia), digested with Not I and cloned into
                     the Not I and Eco RI sites of a modified pT7T3 vector
                     (Pharmacia). Library constructed by Bento Soares and
                     M.Fatima Bonaldo. RNA from normal foreskin melanocytes
                     (FS374) was kindly provided by Dr. Anthony P. Albino."
ORIGIN

Alignment Scores:
Pred. No.:          2.96e-45              Length:          532
Score:              410.00                Matches:         81
Percent Similarity: 100.00%               Conservative:    0
Percent Local Similarity: 100.00%         Mismatches:      0
Query Match:        14                    Indels:          0
                                          Gaps:            0

US-09-648-310-4 (1-81) x N39717 (1-532)

y     1 MetAsnValAspHisGluAlaLeuValGlyThrLeuLysAlaAlaLysArgArgLysIleValThr 60
Qy      1 MetAsnValAspHisGluAlaLeuValGlyLeuValAsnLeuValValGluIleHisArgLeuGlySerLys 20
b    90 ATGAATGTGGATCACGAGGTTAACCTCTTAGTGGAGGAAATTCATCGTTTGGGTTCAAAA 149

Y    21 AsnAlaAspGlyLysLysLysPheValLysSerValLysSerPheValLeuPheArgAspAspLysCysAla 40
b   150 AATGCTGATGGAAAGTTAAGCGTGAAATTTGGGTCCTCTTCCGTGATGATAAATGTGCC 209

Y    41 AsnLeuPheGluGluLeuValGlyValGlyThrLeuLysAlaAlaLysArgArgLysIleValThr 60
b   210 AACCTCTTTGAAGCATTGGTAGGAACCTTAAAGCTGCAAAACGAAGGAAGATTGTAACA 269

Y    61 TyrProGlyLeuLeuLeuGlnLeuGlyValHisAspAspValAspIleIleLeuLeuGln 80

Best Local Similarity:    100.00%        Mismatches:    0
Query Match:              100.00%        Indels:        0
DB:                       13             Gaps:          0

US-09-648-310-4 (1-81) x BX474507 (1-544)

```
Qy    1 MetAsnValAspHisGluValAsnLeuLeuValGluGluIleHisArgLeuGlySerLys  20
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   68 ATGAATGTGGATCACGAGGTTAACCTCTTAGTGGAGGAAATTCATCGTTTGGGTTCAAAA 127

Qy   21 AsnAlaAspGlyLysLeuSerValLysPheGlyValLeuPheArgAspLysCysCysAla  40
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  128 AATGCTGATGGAAAGTTAAGCGTGAAATTTGGGTCCTCTTCCGTGATGATAAATGTGCC 187

Qy   41 AsnLeuPheGluAlaLeuValGlyThrLeuLysAlaAlaLysArgArgLysIleValThr  60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  188 AACCTCTTTGAAGCATTGGTAGGAACTCTTAAAGCTGCAAAACGAAGGAAGATTGTAACA 247

Qy   61 TyrProGlyGluLeuLeuGlnGlyValHisAspAspValAspIleIleLeuLeuGln  80
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  248 TATCCAGGAGAGCTGCTTCTGCAAGGTGTTCATGATGATGTTGACATTATTACTGCAA 307

Qy   81 Asp 81
        |||
Db  308 GAT 310
```

RESULT 6
BU198216
LOCUS   BU198216   567 bp   mRNA   linear   EST 05-SEP-2002
DEFINITION DCBCNC04 DCB Homo sapiens cDNA, mRNA sequence.
ACCESSION   BU198216
VERSION   BU198216.1 GI:22716103
KEYWORDS   EST.
SOURCE   Homo sapiens (human)
  ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 567)
  AUTHORS   Xu,X., Gu,J., Liu,F., Qu,J., Zhao,M., Li,Y., Huang,Q., Zhou,J.,
            Song,H., Gu,Y., Yang,Y., Gao,G., Xiao,H., Li,N., Qian,B., Gao,X.,
            Cheng,Z., Xu,S., Gu,W., Tu,Y., Jia,J., Fu,G., Ren,S., Zhong,M.,
            Lu,G., Cheng,Z. and Han,Z.
  TITLE   Homo sapiens cDNA DCB clones
  JOURNAL   Unpublished (2000)
COMMENT   Contact: Zeguang Han
            Chinese National Human Genome Center at Shanghai
            351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
            201203, P. R. China
            Tel: 86-21-50801919(ex.45)
            Fax: 86-21-50801922
            Email: hanzg@chgc.sh.cn.
FEATURES   Location/Qualifiers
     source   1..567
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /cell_type="dendritic cells"
                     /dev_stage="mature"
                     /lab_host="BM25.8"
                     /clone_lib="DCB"
                     /note="Vector: pTriplEx2; Site_1: sfiIA; Site_2: sfiIB"
ORIGIN

Alignment Scores:
Pred. No.:   3.22e-45   Length: 567
Score: 410.00   Matches: 81
Percent Similarity: 100.00%   Conservative: 0
Best Local Similarity: 100.00%   Mismatches: 0
Query Match: 100.00%   Indels: 0
DB: 13   Gaps: 0

US-09-648-310-4 (1-81) x BU198216 (1-567)

Best Local Similarity: 100.00%   Mismatches: 0
Query Match: 100.00%   Indels: 0
DB: 12   Gaps: 0

US-09-648-310-4 (1-81) x BG944932 (1-542)

```
Qy    1 MetAsnValAspHisGluValAsnLeuLeuValGluGluIleHisArgLeuGlySerLys  20
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  192 ATGAATGTGGATCACGAGGTTAACCTCTTAGTGGAGGAAATTCATCGTTTGGGTTCAAAA 251

Qy   21 AsnAlaAspGlyLysLeuSerValLysPheGlyValLeuPheArgAspLysCysCysAla  40
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  252 AATGCTGATGGAAAGTTAAGCGTGAAATTTGGGTCCTCTTCCGTGATGATAAATGTGCC 311

Qy   41 AsnLeuPheGluAlaLeuValGlyThrLeuLysAlaAlaLysArgArgLysIleValThr  60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  312 AACCTCTTTGAAGCATTGGTAGGAACTCTTAAAGCTGCAAAACGAAGGAAGATTGTAACA 371

Qy   61 TyrProGlyGluLeuLeuGlnGlyValHisAspAspValAspIleIleLeuLeuGln  80
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  372 TATCCAGGAGAGCTGCTTCTGCAAGGTGTTCATGATGATGTTGACATTATTACTGCAA 431

Qy   81 Asp 81
        |||
Db  432 GAT 434
```

RESULT 5
X474507
LOCUS   BX474507   544 bp   mRNA   linear   EST 04-SEP-2003
DEFINITION DKFZp686E19170_r1 686 (synonym: hlcc3) Homo sapiens cDNA clone
            DKFZp686E19170 5', mRNA sequence.
ACCESSION   BX474507
VERSION   BX474507.1 GI:31668714
KEYWORDS   EST.
SOURCE   Homo sapiens (human)
  ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 544)
  AUTHORS   Bloecker,H., Boecher,M., Mewes,H.W., Weil,B., Amid,C., Osanger,A.,
            Fobo,G., Han,M. and Wiemann,S.
  TITLE   EST (Bloecker,H., Boecher,M., Mewes,H.W., Weil,B., Amid,C., et al.)
  JOURNAL   Unpublished (2003)
COMMENT   Contact: MIPS
            MIPS
            Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
            This is the 5, sequence of the clone insert
            Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
            Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
            sequenced by GBF (National Research Centre for Biotechnology Ltd.,
            Braunschweig/Germany) within the cDNA sequencing consortium of the
            German Genome Project.
            No sl sequence available.
            This clone (DKFZp686E19170) is available at the RZPD in Berlin.
            Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
            Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
FEATURES   Location/Qualifiers
     source   1..544
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="DKFZp686E19170"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /clone_lib="686 (synonym: hlcc3)"
                     /note="Vector: pTriplEx2; Site_1: sfiIA; Site_2: sfiIB;
                     cDNA-collection"
ORIGIN

Alignment Scores:
Pred. No.:   3.05e-45   Length: 544
Score: 410.00   Matches: 81
Percent Similarity: 100.00%   Conservative: 0

Right column

```
Alignment Scores:
Pred. No.:
Score:                      3.26e-45        Length:        571
Percent Similarity:         410.00          Matches:       81
Best Local Similarity:      100.00%         Conservative:  0
Query Match:                100.00%         Mismatches:    0
DB:                                         Indels:        0
                                            Gaps:          0

US-09-648-310-4 (1-81) x BE145690 (1-571)

Qy   1 MetAsnValAspHisGluValAsnLeuLeuValGluGluIleHisArgLeuGlySerLys 20
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  24 ATGAATGTGGATCACGAGGTTAACCTCTTAGTGGAGGAAATTCATCGTTTGGGTTCAAAA 83

Qy  21 AsnAlaAspGlyLysLeuSerValLysPheArgAspAspLysCysAla 40
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  84 AATGCTGATGGAAAGTTAAGCGTGAAATTTGGGGTCCTCTTCCGTGATGATAAATGTGCC 143

Qy  41 AsnLeuPheGluAlaLeuValGlyThrLeuLysAlaAlaLysArgArgLysIleValThr 60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 144 AACCTCTTTGAAGCATTGGTAGGAACTCTTAAAGCTGCAAAACGAAGGAAGATTGTAACA 203

Qy  61 TyrProGlyGluLeuLeuGlnGlyValHisAspAspValAspIleIleIleLeuLeuGln 80
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 204 TATCCAGGAGAGCTGCTTCTGCAAGGTGTTCATGATGTTGACATTATTATTACTGCAA 263

Qy  81 Asp 81
         |||
Db 264 GAT 266

RESULT 8
CA335746
LOCUS       CA335746            576 bp    mRNA    linear   EST 04-NOV-2002
DEFINITION  NISC_lu03e04.y1 COGENE 4AR (4EAR) Homo sapiens cDNA clone
            IMAGE:5606886 5', mRNA sequence.
ACCESSION   CA335746
VERSION     CA335746.1  GI:24553844
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 576)
AUTHORS     NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE       National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
JOURNAL     Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            cDNA library Preparation:
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium/LLNL
            DNA Sequencing by: National Institutes of Health Intramural
              Sequencing Center (NISC)
            Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            info@image.llnl.gov
            Plate: LLAM12407 row: J column: 7
            Seq primer: M13RP1 reverse primer (ABI).
FEATURES    Location/Qualifiers
     source 1..576
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone="IMAGE:5606886"
            /tissue_type="anterior rhombomeres 1-4 (4 pooled)"
            /dev_stage="embryo, 4 weeks postconception"
            /lab_host="DH10B"
            /clone_lib="COGENE 4AR (4EAR)"
            /note="Vector: pAMP1; cDNA primed using oligo-dT primer,
            directionally cloned into UDG sites of pAMP1. Size
            selected for insert sizes ranging from 0.3-1.6 kb.
            Normalized to Cot10. Primary library, non-amplified.
```

Left column

```
Qy    1 MetAsnValAspHisGluValAsnLeuLeuValGluGluIleHisArgLeuGlySerLys 20
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  201 ATGAATGTGGATCACGAGGTTAACCTCTTAGTGGAGGAAATTCATCGTTTGGGTTCAAAA 260

Qy   21 AsnAlaAspGlyLysLeuSerValLysPheArgAspAspLysCysAla 40
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  261 AACGCTGATGGAAAGTTAAGCGTGAAATTTGGGGTCCTCTTCCGTGATGATAAATGTGCC 320

Qy   41 AsnLeuPheGluAlaLeuValGlyThrLeuLysAlaAlaLysArgArgLysIleValThr 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  321 AACCTCTTTGAAGCATTGGTAGGAACTCTTAAAGCTGCAAAACGAAGGAAGATTGTAACA 380

Qy   61 TyrProGlyGluLeuLeuGlnGlyValHisAspAspValAspIleIleIleLeuLeuGln 80
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  381 TATCCAGGAGAGCTGCTTCTGCAAGGTGTTCATGATGTTGACATTATTATTACTGCAA 440

Qy   81 Asp 81
          |||
Db  441 GAT 443

SULT 7
145690
CUS
FINITION  BE145690            571 bp    mRNA    linear   EST 21-JUN-2000
          IL0-HT0205-131199-143-b09 HT0205 Homo sapiens cDNA, mRNA sequence.
CESSION   BE145690
RSION     BE145690.1  GI:8608414
YWORDS    EST.
URCE      Homo sapiens (human)
ORGANISM  Homo sapiens
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
FERENCE 1 (bases 1 to 571)
AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
           Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
           Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
           Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
           O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
           Simpson,A.J.
TITLE     Shotgun sequencing of the human transcriptome with ORF expressed
           sequence tags
JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE   20202663
PUBMED    10737800
OMMENT    Contact: Simpson A.J.G.
           Laboratory of Cancer Genetics
           Ludwig Institute for Cancer Research
           Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
           Brazil
           Tel: +55-11-2704922
           Fax: +55-11-2707001
           Email: asimpson@ludwig.org.br
           This sequence was derived from the FAPESP/LICR Human Cancer Genome
           Project. This entry can be seen in the following URL
           (http://www.ludwig.org.br/scripts/gethtm12.pl?t1=&t2=IL0-HT0205-131
           199-143-b09&t3=1999-11-13&t4=1)
           Seq primer: puc 18 forward
           High quality sequence start: 26
           High quality sequence stop: 571
EATURES   Location/Qualifiers
    source 1..571
           /organism="Homo sapiens"
           /mol_type="mRNA"
           /db_xref="taxon:9606"
           /dev_stage="Adult"
           /clone_lib="HT0205"
           /note="Organ: head; neck; Vector: puc18; Site 1: SmaI;
           Site_2: SmaI; A mini-library was made by cloning products
           derived from ORESTES PCR (U.S. Letters Patent application
           No. 196,716 - Ludwig Institute for Cancer Research).
           profiles into the pUC 18 vector. Reverse transcription of
           tissue mRNA and cDNA amplification were performed under
           low stringency conditions."
RIGIN
```

ORIGIN

/clone_lib="TP"
/note="Vector: pTriplEx2; Site_1: sfiIA; Site_2: sfiIB"

Alignment Scores:
Pred. No.:
Score: 3.47e-45   Length: 599
Percent Similarity: 410.00   Matches: 81
Best Local Similarity: 100.00%   Conservative: 0
Query Match: 100.00%   Mismatches: 0
DB: 100.00%   Indels: 0
                                            Gaps: 0

US-09-648-310-4 (1-81) x AV753954 (1-599)

Qy   1 MetAsnValAspHisGluValAsnLeuValGluGluIleHisArgLeuGlySerLys 20
Db 172 ATGAATGTGGATCACGAGGTTAACCTCTTAGTGGAGGAAATTCATCGTTGGGTTCAAAA 231

Qy   21 AsnAlaAspGlyLysLeuSerValLysPheGlyValLeuPheArgAspAspLysCysAla 40
Db 232 AATGCTGATGGAAAGTTAAGCGTGAAATTTGGGGTCCTCTTCCGTGATGAAAATGTGCC 291

Qy   41 AsnLeuPheGluAlaLeuValGlyThrLeuLysAlaAlaLysArgLysIleValThr 60
Db 292 AACCCTCTTTGAAGCATTGGTAGGAACTCTTAAAGCTGCAAAACGAAGGAAGATTGTAACA 351

Qy   61 TyrProGlyGluLeuLeuLeuGlnGlyValHisAspAspValAspIleIleLeuLeuGln 80
Db 352 TATCCAGGAGAGCTGCTTCTGCAAGGTGTTCATGATGATGTTGACATTATTATTACTGCAA 411

Qy   81 Asp 81
Db 412 GAT 414

RESULT 10
CB162901
LOCUS   CB162901   599 bp   mRNA   linear   EST 30-JAN-2003
DEFINITION K-EST0223598 L18POOLln1 Homo sapiens cDNA clone L18POOLln1-41-G10
            5', mRNA sequence.
ACCESSION   CB162901
VERSION   CB162901.1   GI:28149027
KEYWORDS   EST.
SOURCE   Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 599)
AUTHORS   Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
            Kim,Y.S., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
            Kim,Y.S.
TITLE   21C Frontier Korean EST Project 2001
JOURNAL   Unpublished (2002)
COMMENT   Contact: Kim YS
            Genome Research Center
            Korea Research Institute of Bioscience & Biotechnology
            52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
            Tel: +82-42-860-4470
            Fax: +82-42-860-4409
            Email: yongsung@mail.kribb.re.kr
            Plate: 41 row: G column: 10
            High quality sequence stop: 599.
FEATURES   Location/Qualifiers
     source   1..599
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="L18POOLln1-41-G10"
                     /cell_line="SNU-354+Cho-CK+Choi-CK+HLK-3"
                     /lab_host="Top10F'"
                     /clone_lib="L18POOLln1"
                     /note="Organ: Liver; Vector: pT7T3-Pac; Site_1: EcoRI;
                     Site_2: NotI; The library was contributed by the Soares
                     laboratory and it was constructed as described by Bonaldo,

ORIGIN

Library constructed by M. Lovett. For more information on
this library, please contact R. Tidwell (Washington
University) or visit the COGENE website at
http://hg.wustl.edu/COGENE/."

Alignment Scores:
Pred. No.:
Score: 3.29e-45   Length: 576
Percent Similarity: 410.00   Matches: 81
Score: 100.00%   Conservative: 0
Best Local Similarity: 100.00%   Mismatches: 0
Query Match: 100.00%   Indels: 0
DB: Gaps: 0

S-09-648-310-4 (1-81) x CA335746 (1-576)

Qy   1 MetAsnValAspHisGluValAsnLeuValGluGluIleHisArgLeuGlySerLys 20
Db   82 ATGAATGTGGATCACGAGGTTAACCTCTTAGTGGAGGAAATTCATCGTTGGGTTCAAAA 141

Qy   21 AsnAlaAspGlyLysLeuSerValLysPheGlyValLeuPheArgAspAspLysCysAla 40
Db 142 AATGCTGATGGAAAGTTAAGCGTGAAATTTGGGGTCCTCTTCCGTGATGAAAATGTGCC 201

Qy   41 AsnLeuPheGluAlaLeuValGlyThrLeuLysAlaAlaLysArgLysIleValThr 60
Db 202 AACCCTCTTTGAAGCATTGGTAGGAACTCTTAAAGCTGCAAAACGAAGGAAGATTGTAACA 261

Qy   61 TyrProGlyGluLeuLeuLeuGlnGlyValHisAspAspValAspIleIleLeuLeuGln 80
Db 262 TATCCAGGAGAGCTGCTTCTGCAAGGTGTTCATGATGATGTTGACATTATTATTACTGCAA 321

Qy   81 Asp 81
Db 322 GAT 324

RESULT 9
AV753954
LOCUS   AV753954   599 bp   mRNA   linear   EST 19-OCT-2000
DEFINITION AV753954 TP Homo sapiens cDNA clone TPAASB04 5', mRNA sequence.
ACCESSION   AV753954
VERSION   AV753954.1   GI:10911802
KEYWORDS   EST.
SOURCE   Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 599)
AUTHORS   Song,H., Peng,Y., Gu,Y., Yang,Y., Gao,G., Xiao,H., Xu,X., Li,N.,
            Qian,B., Liu,F., Qu,J., Gao,X., Cheng,Z., Xu,Z., Zeng,L., Xu,S.,
            Gu,W., Tu,Y., Jia,J., Fu,G., Ren,S., Zhong,M., Lu,G., Hu,R., Ye,M.,
            Zhang,Q., Han,Z., Chen,Z. and Chen,J.
TITLE   Homo sapiens TP library cDNA clones
JOURNAL   Unpublished (2000)
COMMENT   Contact: Qinghua Zhang
            Shanghai Institute of Endocrinology, Rui-Jin Hospital
            197 Rui-Jin II Road, Shanghai 200025, P. R. China
            Tel: 86-21-64370045(ex.663332)
            Fax: 86-21-64743206
            Email: mbshi@ms.stn.sh.cn
            This clone is available at Shanghai Hematology Institute in
            Shanghai.
            Chinese National Human Genome Center at Shanghai
            351 Guo Shouing Road, Zhangjiang Hi-Tech Park, Pudong.
FEATURES   Location/Qualifiers
     source   1..599
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="TPAASB04"
                     /tissue_type="pituitary tumor"
                     /dev_stage="Adult"
                     /lab_host="BM25.8"

```
RESULT 11
AV714738
LOCUS       AV714738
DEFINITION  AV714738 DCB Homo sapiens cDNA clone DCBBRC03 5', mRNA sequence.
ACCESSION   AV714738
VERSION     AV714738.1  GI:10796255
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 601)
AUTHORS     Xu,X., Gu,J., Liu,F., Qu,J., Zhao,M., Li,Y., Huang,Q., Zhou,J.,
            Song,H., Gu,Y., Yang,Y., Gao,G., Xiao,H., Li,N., Qian,B., Gao,X.,
            Cheng,Z., Xu,S., Gu,W., Tu,Y., Jia,J., Fu,G., Ren,S., Zhong,M.,
            Lu,G., Cheng,Z. and Han,Z.
TITLE       Homo sapiens cDNA DCB clones
JOURNAL     Unpublished (2000)
COMMENT     Contact: Zeguang Han
            Chinese National Human Genome Center at Shanghai
            351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
            201203, P. R. China
            Tel: 86-21-50801919(ex.45)
            Fax: 86-21-50801922
            Email: hanz@chgc.sh.cn
            This clone is available at CHGC in Shanghai.
FEATURES             Location/Qualifiers
     source          1..601
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="DCBBRC03"
                     /cell_type="dendritic cells"
                     /dev_stage="mature"
                     /lab_host="BM25.8"
                     /clone_lib="DCB"
                     /note="Vector: pTriplEx2; Site_1: sfiIA; Site_2: sfiIB"
RIGIN
```

```
Alignment Scores:
Pred. No.:             3.47e-45      Length:         599
Score:                 410.00        Matches:        81
Percent Similarity:    100.00%       Conservative:   0
Best Local Similarity: 100.00%       Mismatches:     0
Query Match:           100.00%       Indels:         0
:                      14            Gaps:           0

-09-648-310-4 (1-81) x CB162901 (1-599)

     1 MetAsnValAspHisGluValAsnLeuLeuValGluIleHisArgLeuGlySerLys 20
   146 ATGAATGTGGATCACGAGGTTAACCTCTTAGTGGAGGAATTCATCGTTGGGTTCAAAA 205

    21 AsnAlaAspGlyLysLeuSerValLysPheGlyValLeuPheArgAspAspLysCysAla 40
   206 AATGCTGATGGAAAGTTAAGCGTGAAATTGGGCGTCCTCTTCCGTGATGATAAATGTGCC 265

    41 AsnLeuPheGluAlaLeuValGlyThrLeuLysAlaAlaLysArgArgLysIleValThr 60
   266 AACCTCTTTGAAGCATTGGTAGGAACTCTTAAAGCTGCAAAACGAAGGAAGATTGTAACA 325

    61 TyrProGlyGluLeuLeuLeuGlnGlyValHisAspAspValAspIleIleLeuLeuGln 80
   326 TATCCAGGAGAGCTGCTTCTGCAAGGTGTTCATGATGATGTTGACATTATATTACTGCAA 385

    81 Asp 81
   386 GAT 388
```

```
Alignment Scores:
Pred. No.:             3.49e-45      Length:         601
Score:                 410.00        Matches:        81
Percent Similarity:    100.00%       Conservative:   0
Best Local Similarity: 100.00%       Mismatches:     0
Query Match:           100.00%       Indels:         0
DB:                    9             Gaps:           0

US-09-648-310-4 (1-81) x AV714738 (1-601)

Qy    1 MetAsnValAspHisGluValAsnLeuLeuValGluIleHisArgLeuGlySerLys 20
Db   96 ATGAATGTGGATCACGAGGTTAACCTCTTAGTGGAGGAATTCATCGTTGGGTTCAAAA 155

Qy   21 AsnAlaAspGlyLysLeuSerValLysPheGlyValLeuPheArgAspAspLysCysAla 40
Db  156 AATGCTGATGGAAAGTTAAGCGTGAAATTGGGCGTCCTCTTCCGTGATGATAAATGTGCC 215

Qy   41 AsnLeuPheGluAlaLeuValGlyThrLeuLysAlaAlaLysArgArgLysIleValThr 60
Db  216 AACCTCTTTGAAGCATTGGTAGGAACTCTTAAAGCTGCAAAACGAAGGAAGATTGTAACA 275

Qy   61 TyrProGlyGluLeuLeuLeuGlnGlyValHisAspAspValAspIleIleLeuLeuGln 80
Db  276 TATCCAGGAGAGCTGCTTCTGCAAGGTGTTCATGATGATGTTGACATTATATTACTGCAA 335

Qy   81 Asp 81
Db  336 GAT 338
```

```
RESULT 12
AV716706
LOCUS       AV716706
DEFINITION  AV716706 DCB Homo sapiens cDNA clone DCBCFB05 5', mRNA sequence.
ACCESSION   AV716706
VERSION     AV716706.1  GI:10813858
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 601)
AUTHORS     Xu,X., Gu,J., Liu,F., Qu,J., Zhao,M., Li,Y., Huang,Q., Zhou,J.,
            Song,H., Gu,Y., Yang,Y., Gao,G., Xiao,H., Li,N., Qian,B., Gao,X.,
            Cheng,Z., Xu,S., Gu,W., Tu,Y., Jia,J., Fu,G., Ren,S., Zhong,M.,
            Lu,G., Cheng,Z. and Han,Z.
TITLE       Homo sapiens cDNA DCB clones
JOURNAL     Unpublished (2000)
COMMENT     Contact: Zeguang Han
            Chinese National Human Genome Center at Shanghai
            351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
            201203, P. R. China
            Tel: 86-21-50801919(ex.45)
            Fax: 86-21-50801922
            Email: hanz@chgc.sh.cn
            This clone is available at CHGC in Shanghai.
FEATURES             Location/Qualifiers
     source          1..601
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="DCBCFB05"
                     /cell_type="dendritic cells"
                     /dev_stage="mature"
                     /lab_host="BM25.8"
                     /clone_lib="DCB"
                     /note="Vector: pTriplEx2; Site_1: sfiIA; Site_2: sfiIB"
ORIGIN

Alignment Scores:
Pred. No.:             3.49e-45      Length:         601
Score:                 410.00        Matches:        81
```

At top (left column header):

```
IGIN
            M.F., Lennon, G. and Soares, M.B. (1996), Genome Research
            6(9):791-806. RNA was prepared from harvested cell
            culture."
```

Percent Similarity:　　　100.00%　　　Conservative:　　　0
Best Local Similarity:　100.00%　　　Mismatches:　　　　0
Query Match:　　　　　　　100.00%　　　Indels:　　　　　　0
DB:　　　　　　　　　　　　　　9　　　　　　Gaps:　　　　　　　0

US-09-648-310-4 (1-81) x AV716706 (1-601)

Qy　　　1 MetAsnValAspHisGluValAsnLeuLeuValGluIleHisArgLeuGlySerLys 20
　　　　　|||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db　　193 ATGAATGTGGATCACGAGGTTAACCTCTTAGTGGAGGAAATTCATCGTTTGGGTTCAAAA 252

Qy　　21 AsnAlaAspGlyLysLeuSerValLysPheGlyValLeuPheArgAspAspLysCysAla 40
　　　　　|||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db　　253 AATGCTGATGGAAAGTTAAGCGTGAAATTTGGGTCCTCTTCCGTGATGATAAATGTGCC 312

Qy　　41 AsnLeuPheGluAlaLeuValGlyThrLeuLysAlaAlaLysArgArgLysIleValThr 60
　　　　　|||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db　　313 AACCTCTTTGAAGCTTTGGTAGGAACTCTTAAAGCTGCAAAACGAAGGAAGATTGTAACA 372

Qy　　61 TyrProGlyGluLeuLeuGlnGlyValHisAspAspValAspIleIleLeuLeuGln 80
　　　　　|||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db　　373 TATCCAGGAGAGCTGCTTCTGCAAGGTGTTCATGATGATGTTGACATTATTACTGCAA 432

Qy　　81 Asp 81
　　　　　|||
Db　　433 GAT 435

RESULT 13
BU581717
LOCUS　　　BU581717 in40cl1.y1 Human Fetal Pancreas 1B　mRNA　linear　EST 17-SEP-2002
DEFINITION　TR:081881 081881 HYPOTHETICAL 18.3 KD PROTEIN. ; mRNA sequence.
ACCESSION　BU581717
VERSION　　BU581717.1　GI:23066977
KEYWORDS　　EST.
SOURCE　　　Homo sapiens (human)
ORGANISM　　Homo sapiens
　　　　　　　Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
　　　　　　　Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE　1　(bases 1 to 604)
AUTHORS　　Melton,D., Brown,J., Kenty,G., Permutt,A., Lee,C., Kaestner,K.,
　　　　　　　Lemishka,I., Scearce,M., Brestelli,J., Gradwohl,G., Clifton,S.,
　　　　　　　Hillier,L., Marra,M., Pape,D., Wylie,T., Martin,J., Blistain,A.,
　　　　　　　Schmitt,A., Theising,B., Ritter,E., Ronko,I., Bennett,J.,
　　　　　　　Cardenas,M., Gibbons,M., McCann,R., Cole,R., Tsagareishvili,R.,
　　　　　　　Williams,T., Jackson,Y. and Bowers,Y.
TITLE　　　　Endocrine Pancreas Consortium
JOURNAL　　Unpublished (2000)
COMMENT　　Other_ESTs: in40cl1.x1
　　　　　　　Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
　　　　　　　Endocrine Pancreas Consortium
　　　　　　　Harvard University, Howard Hughes Medical Institute
　　　　　　　Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
　　　　　　　MA 02138
　　　　　　　Tel: 617-495-1812
　　　　　　　Fax: 617-495-8557
　　　　　　　Email: dmelton@biohp.harvard.edu
　　　　　　　Library was constructed by Dr. Hiroshi Inoue DNA sequencing by:
　　　　　　　Washington University Genome Sequencing Center For information on
　　　　　　　obtaining a clone please contact: Dr. Hiroshi Inoue
　　　　　　　(hinoue@imgate.wustl.edu)
　　　　　　　Seq primer: -40RP from Gibco
　　　　　　　High quality sequence stop: 450.
FEATURES　　　　　　　Location/Qualifiers
　　source　　　　　　1..604
　　　　　　　　　　　　/organism="Homo sapiens"
　　　　　　　　　　　　/mol_type="mRNA"
　　　　　　　　　　　　/db_xref="taxon:9606"
　　　　　　　　　　　　/tissue_type="Fetal Pancreas (4 Pooled Donors, 18 - 20
　　　　　　　　　　　　weeks, Stratagene #738023)"
　　　　　　　　　　　　/dev_stage="Fetal Pancreas"
　　　　　　　　　　　　/clone_lib="Human Fetal Pancreas 1B"
　　　　　　　　　　　　/note="Vector: pBluescript SK(-); Site_1: NotI; Site_2:

XhoI; cDNA made by oligo-dT priming. Size-selected on
agarose gel. Average insert size ~1kb. 5' XhoI site was
destroyed after directional cloning. Amplified once.
Contact information: Hiroshi Inoue, MD, Metabolism Div.
(Alan Permutt Lab), Washington University School of
Medicine, Box 8127, 660 South Euclid Ave., St. Louis, MO
63110 E-mail: hinoue@imgate.wustl.edu Tel: 314-362-1916,
Fax:314-747-2692."

ORIGIN

Alignment Scores:
Pred. No.:　　　　　　　3.51e-45　　　　Length:　　　　　　604
Score:　　　　　　　　　410.00　　　　　Matches:　　　　　81
Percent Similarity:　100.00%　　　Conservative:　　0
Best Local Similarity:100.00%　　　Mismatches:　　　0
Query Match:　　　　　100.00%　　　Indels:　　　　　0
DB:　　　　　　　　　　13　　　　　　　Gaps:　　　　　　0

US-09-648-310-4 (1-81) x BU581717 (1-604)

Qy　　1 MetAsnValAspHisGluValAsnLeuLeuValGluIleHisArgLeuGlySerLys 20
　　　　|||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db　　45 ATGAATGTGGATCACGAGGTTAACCTCTTAGTGGAGGAAATTCATCGTTTGGGTTCAAAA 104

Qy　　21 AsnAlaAspGlyLysLeuSerValLysPheGlyValLeuPheArgAspAspLysCysAla 40
　　　　|||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db　105 AATGCTGATGGAAAGTTAAGCGTGAAATTTGGGTCCTCTTCCGTGATGATAAATGTGCC 164

Qy　　41 AsnLeuPheGluAlaLeuValGlyThrLeuLysAlaAlaLysArgArgLysIleValThr 60
　　　　|||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db　165 AACCTCTTTGAAGCTTTGGTAGGAACTCTTAAAGCTGCAAAACGAAGGAAGATTGTAACA 224

Qy　　61 TyrProGlyGluLeuLeuGlnGlyValHisAspAspValAspIleIleLeuLeuGln 80
　　　　|||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db　225 TATCCAGGAGAGCTGCTTCTGCAAGGTGTTCATGATGATGTTGACATTATTACTGCAA 284

Qy　　81 Asp 81
　　　　|||
Db　285 GAT 287

RESULT 14
AI133406
LOCUS　　　AI133406　　　　　　　　　　　634 bp　　mRNA　　　linear　EST 11-NOV-1999
DEFINITION　HA2013 Human fetal liver cDNA library Homo sapiens cDNA, mRNA
　　　　　　　sequence.
ACCESSION　AI133406
VERSION　　AI133406.1　GI:6360722
KEYWORDS　　EST.
SOURCE　　　Homo sapiens (human)
ORGANISM　　Homo sapiens
　　　　　　　Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
　　　　　　　Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE　1　(bases 1 to 634)
AUTHORS　　Yu,Y., Zhang,C., Luo,L., Ouyang,S., Li,W., Wu,J., Zhou,S., Liu,M.
　　　　　　　and He,F.
TITLE　　　　Expression profile analysis of a human fetal liver cDNA library
JOURNAL　　Unpublished (1998)
COMMENT　　Contact: Yongtao Yu
　　　　　　　Department of Hematology,
　　　　　　　Beijing Institute of Radiation Medicine
　　　　　　　27 Taiping Road, Beijing 100850, P.R.China
　　　　　　　Tel: 0086-10-68159479
　　　　　　　Fax: 0086-10-68214653
　　　　　　　Email: yyt48@yahoo.com.
FEATURES　　　　　　　Location/Qualifiers
　　source　　　　　　1..634
　　　　　　　　　　　　/organism="Homo sapiens"
　　　　　　　　　　　　/mol_type="mRNA"
　　　　　　　　　　　　/db_xref="taxon:9606"
　　　　　　　　　　　　/tissue_type="liver"
　　　　　　　　　　　　/dev_stage="fetal"
　　　　　　　　　　　　/lab_host="MC1061/P3"
　　　　　　　　　　　　/clone_lib="Human fetal liver cDNA library"

C, or G and N = A, C, G, or T). Average insert size 1.85
kb (range 1.0-4.0 kb). 15/15 colonies contained inserts
by PCR. This library was enriched for full-length clones
and was constructed by Clontech Laboratories (Palo Alto,
CA). Note: this is a NIH_MGC Library."

ORIGIN

```
Alignment Scores:
Pred. No.:                                Length:           641
Score:                 3.8e-45            Matches:           81
Percent Similarity:    410.00             Conservative:       0
Best Local Similarity: 100.00%            Mismatches:         0
Query Match:           100.00%            Indels:             0
DB:                    12                 Gaps:               0
```

US-09-648-310-4 (1-81) x BG563686 (1-641)

```
Qy   1 MetAsnValAspHisGluValAsnLeuValGluGluIleHisArgLeuGlySerLys  20
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 202 ATGAATGTGGATCACGAGGTTAACCTCTTAGTGGAGGAAATTCATCGTTTGGGTTCAAAA 261

Qy  21 AsnAlaAspGlyLysLeuSerValLysPheGlyValLeuPheArgAspAspLysCysAla  40
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 262 AATGCTGATGGAAAGTTAAGCGTGAAATTTGGGGTCCTTCTTCCGTGATGATAAATGTGCC 321

Qy  41 AsnLeuPheGluAlaLeuValGlyThrLeuArgArgLysAlaAlaLysArgLysIleValThr  60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 322 AACCTCTTTGAAGCATTGGTAGGAACTCTTAAAGCTGCAAAACGAAGGAAGATTGTAACA 381

Qy  61 TyrProGlyLeuLeuLeuGlnGlyLeuValHisAspAspValAspIleIleLeuLeuGln  80
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 382 TATCCAGGAGAGCTGCTTCTGCAAGGTGTTCATGATGATGTTGACATTATTACTGCAA 441

Qy  81 Asp  81
         |||
Db 442 GAT 444
```

Search completed: March 17, 2004, 18:24:08
Job time : 1970 secs

ORIGIN
                /note="Vector: pCDNAl"

```
Alignment Scores:
Pred. No.:                                Length:           634
Score:                 3.75e-45           Matches:           81
Percent Similarity:    410.00             Conservative:       0
Best Local Similarity: 100.00%            Mismatches:         0
Query Match:           100.00%            Indels:             0
DB:                    9                  Gaps:               0
```

US-09-648-310-4 (1-81) x AI133406 (1-634)

```
Qy   1 MetAsnValAspHisGluValAsnLeuValGluGluIleHisArgLeuGlySerLys  20
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 136 ATGAATGTGGATCACGAGGTTAACCTCTTAGTGGAGGAAATTCATCGTTTGGGTTCAAAA 195

Qy  21 AsnAlaAspGlyLysLeuSerValLysPheGlyValLeuPheArgAspAspLysCysAla  40
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 196 AATGCTGATGGAAAGTTAAGCGTGAAATTTGGGGTCCTTCTTCCGTGATGATAAATGTGCC 255

Qy  41 AsnLeuPheGluAlaLeuValGlyThrLeuArgArgLysAlaAlaLysArgLysIleValGln  60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 256 AACCTCTTTGAAGCATTGGTAGGAACTCTTAAAGCTGCAAAACGAAGGAAGATTGTAACA 315

Qy  61 TyrProGlyLeuLeuLeuGlnGlyLeuValHisAspAspValAspIleIleLeuLeuGln  80
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 316 TATCCAGGAGAGCTGCTTCTGCAAGGTGTTCATGATGATGTTGACATTATTACTGCAA 375

Qy  81 Asp  81
         |||
Db 376 GAT 378
```

RESULT 15
BG563686
LOCUS   BG563686   60258452F1 NIH_MGC_76 Homo sapiens cDNA clone IMAGE:4712125 5',
                       641 bp   mRNA   linear   EST 10-APR-2001
            mRNA sequence.
ACCESSION   BG563686
VERSION   BG563686.1 GI:13571338
KEYWORDS   EST.
SOURCE   Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 641)
AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
TITLE   National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL   Unpublished (199)
COMMENT   Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: CLONETECH Laboratories, Inc.
            cDNA Library Preparation: CLONETECH Laboratories, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM1553 row: h column: 14
            High quality sequence stop: 641.
FEATURES   Location/Qualifiers
     source   1..641
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4712125"
                     /lab_host="DH10B (T1 phage-resistant)"
                     /clone_lib="NIH_MGC_76"
                     /note="Organ: liver; Vector: pDNR-LIB (Clontech); Site 1:
                     SfiI (ggccgcctcggcc); Site 2: SfiI (ggccattatggcc); 5' and
                     3' adaptors were used in cloning as follows: 5' adaptor
                     sequence: 5'-CACGGCCATTATGCC-3' and 3' adaptor sequence:
                     5'-ATTCTAGAGGCCGAGGCGGCCGACATG-dT(30)BN-3' (where B = A,